Some Studies in Radio Broadcast Transmission[1]

By RALPH BOWN, DeLOSS K. MARTIN
and RALPH K. POTTER

SYNOPSIS: The paper is based on radio transmission tests from station 2XB in New York City to two outlying field stations. It is a detailed study of fading and distortion of radio signals under night time conditions in a particular region which may or may not be typical.

Night time fading tests using constant single frequencies and bands of frequencies in which the receiving observations were recorded by oscillograph show that the fading is selective. By selective fading it is meant that different frequencies do not fade together. From the regularity of the frequency relation between the frequencies which fade together it is concluded that the selective fading is caused by wave interference. The signals appear to reach the receiving point by at least two paths of different lengths. The paths change slowly with reference to each other so that at different times the component waves add or neutralize, going through these conditions progressively. The two major paths by which the interfering waves travel are calculated to have a difference in length of the order of 135 kilometers for the conditions of the tests. Since this difference is greater than the distance directly from transmitter to receiver it is assumed that one path at least must follow a circuitous route, probably reaching upward through higher atmospheric regions. Various theories to explain this are briefly reviewed.

The territory about one of the receiving test stations in Connecticut is found under day time conditions to be the seat of a gigantic fixed wave interference or diffraction pattern caused in part by the shadowing of a group of high buildings in New York City. The influence of this pattern on night time fading is discussed. It is considered a contributing but not the controlling effect.

Tests using transmission from an ordinary type of broadcasting transmitter show that such transmitters have a dynamic frequency instability or frequency modulation combined with the amplitude modulation. At night the wave interference effects which produce selective fading result in distortion of the signals when frequency modulation is present. It is shown that stabilizing the transmitter frequency eliminates this distortion. A theory explaining the action is given. The distortions predicted by the theory check with the actual distortions observed.

A discussion of ordinary modulated carrier transmission, carrier suppression, and single side band transmission is given in relation to selective fading It is shown that the use of a carrier suppression system should reduce fading.

ONE of the factors which must be given increasing attention, if the technique of radio telephone broadcasting is to consolidate and continue its remarkable progress, is the mechanism of the transmission of radio signals through space. In many receiving situations the largest apparent defects present in the reproduced signal are those suffered not in the terminal apparatus but in transit through space, and in these cases better methods of utilizing the transmitting medium must precede any major betterment in overall results. In the present paper we are reporting some investigations in this field of radio transmission which have uncovered a number of interesting facts and have led to at least one conclusion which is of practical utility.

[1] Presented before the Institute of Radio Engineers, New York, Nov. 4, 1925

Night time transmission, which is the usual case in broadcasting, is in many places commonly marred by fading and sometimes by actual distortion of signals. Often these occur in certain areas not more distant from the transmitting station than other areas which enjoy freedom from such annoyance. Selecting a particular instance of these difficulties in an area near New York City which, in so far as can be judged at present, is probably a typical instance, we have subjected it to an intensive experimental study to determine what is the inherent nature of the troubles and if possible how they may be alleviated. In doing this it has been necessary to employ novel forms of tests especially fitted to bring out in a concrete way the phenomena being investigated.

To provide a suitable background for the subject we have started our discussion below with a brief recital of some of the things which a transmission medium is called upon to do. Following this we have described our tests, pointing out in what ways the existing media seem to fall short of doing these things and offering certain speculations as to the reasons for the shortcomings. In conclusion we have analyzed some practical problems in the light of this work.

Fundamental Considerations

As the radio art has progressed from spark telegraphy into continuous wave telegraphy and into high quality radio telephone broadcasting, increasing demands have been made on the transmission medium to deliver at the receiving point a true sample of what was put into it at the transmitting station. The requirements have grown in rigor because in telegraphy the end has been to develop increased reliability of communication at longer ranges and in telephony the medium is called upon to transmit a highly complex form of intelligence.

Of the requirements placed on the transmission medium by modern uses, those imposed by telephony are far more exacting than those for telegraphy. In telegraphy a single frequency, or at most a narrow band of frequencies sent out intermittently in accordance with a dot and dash code must reach the receiving station in such shape that it may be converted into audible sound for aural interpretation or into current pulses for the operation of relays or recording instruments. Leaving aside noise, the principal requirement is a sufficient freedom from fading so that signals can be interpreted or recorded without interruption. In radio telephony, as at present practiced in broadcasting, there is transmitted a modulated high-frequency wave comprising a relatively wide band of frequencies, usually at least 10 kilocycles.

Such a modulated high-frequency wave drawn out in the familiar graphical representation is a comparatively simple-looking thing, but analyzed into its elements and studied in detail it is revealed as being an intricate fabric of elemental waves so interwoven with each other that no one of them can be disturbed without changing in some degree the complexion of the whole. For perfect results the whole band must arrive at the receiver with an amplitude continuously proportional to that leaving the transmitter, or the inflections or expression of the speech or music will not be correctly reproduced. All the component frequencies within the band must be unchanged in their relative amplitudes lest the character of the sounds be altered. Even the relative phase relations of the various frequencies must be preserved or, as will be shown later, the interaction of the two side bands in the receiving detector will result in the partial loss of some of the frequency components.

It is not long since the time when radio was supposed to be the perfect medium for voice transmission it being presumed that since the ether of space (if there be such a thing) was substantially perfect in its electrical characteristics it must transmit frequency bands carrying telephone channels without distortion of any kind. This may be true theoretically of a pure ether but in fact, the ether used for radio communication is filled with a number of things ranging from gaseous ions down to the solid bed rock of the earth. It is rather to be expected that these will affect the progress of electromagnetic waves and we know from experience that they do. Diurnal variations of attenuation, fading, directional changes, dead spots and the like are already well known phenomena resulting from the complexity of our transmission media, although no entirely adequate explanations of their causes have been certainly established. One of the most recent manifestations of the effects of irregularities in transmission through space is in the distortion of the quality of telephone signals. This was perhaps first noticed in the use of short waves for broadcasting it being found that frequently the transmission was so distorted that after detection the signals such as speech and music were in severe cases almost unrecognizable.

Preliminary Investigations

For some time after quality distortion was recognized as a characteristic of existing short wave transmissions, it was thought that for the lower broadcasting frequencies at least, it was present only at night and at relatively very great distances from the transmitter. However,

careful observations demonstrated that there were points relatively near New York City where quality distortion from several broadcasting stations in the city was marked at night and in at least one case was detectable even in daytime. When station 2XB the Bell Telephone Laboratories' experimental station at 463 West Street, New York City, was used to transmit test signals, it was found that quality distortion could be observed in northern Westchester county and in southern Connecticut at distances of about 30 to 50 miles from the transmitter. Fading was also pronounced and it was noted as a significant fact that distortion was always accompanied by some fading although the reverse was not consistently true. In the course of these trials it was noticed that at a particular point near New Canaan, Connecticut, signals from 2XB were much weaker and more distorted than signals from 2XY, the experimental station of the American Telephone and Telegraph Company at 24 Walker Street, New York, even though the transmitter at 2XB was about ten times more powerful. Daylight field strength measurements at this point showed that the field strength of 2XB was only one-third that of 2XY. This led to the rather startling conclusion that there is a ratio of 100 to 1 in the power efficiency of transmission to that particular receiving point from these two transmitting stations in New York which are only about one mile apart.

In order to throw some light on this state of affairs a field strength survey was made by G. D. Gillett which resulted in the field strength contour map [1] here reproduced in Fig. 1. The contours on this map show that there is a series of long nearly parallel hills and valleys of field strength which, extrapolated, would converge in lower Manhattan and which extend out to the northeast as far as it was thought worth while to follow them. There has occurred to us no better explanation of this hitherto uncharted form of field strength distribution than that it is a gigantic wave interference pattern. A detailed discussion of this theory is given in another section of this paper.

The fixed pattern shown by Fig. 1 is definitely present only in the daytime but that it is fixed is attested by the fact that a second survey made about a year later checks with the original one quite closely. At night fading is pronounced in the area covered by the pattern and it is apparent that some other factors must enter. As a result of an endeavor to check up the pattern at night it was discovered that

[1] This map was prepared by Mr. Gillett using the methods discussed in a paper "Distribution of Radio Waves from Broadcasting Stations Over City Districts," by Ralph Bown and G. D. Gillett, I. R. E. Proc., Vol. 12, No. 4, p. 395—August, 1924.

quality distortion was, in general, most evident at places which were, by day, in the valleys of the field strength diagram and a point in one of these valleys near Stamford, Connecticut, was selected for the establishment of a temporary field test station. The interior of this station, which was in the empty hay-mow of a barn, is illustrated by the

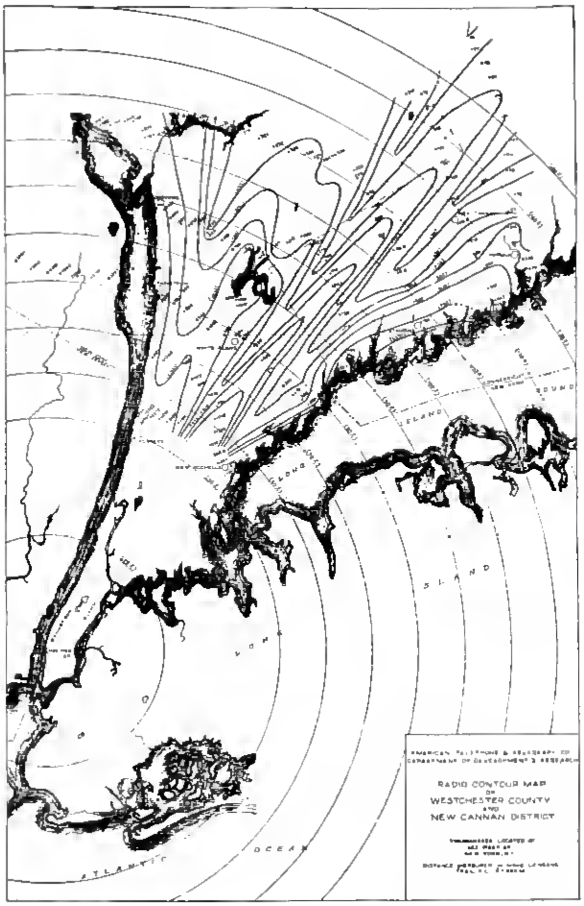

Fig. 1—Radio contour map showing wave interference pattern

photograph, Fig. 2. At this place apparatus was set up to enable a study of the nature of the distortion in signals from 2XB. Many of the records discussed in succeeding paragraphs were taken at this Stamford field station. Others were taken near Riverhead, Long Island, which was also found to be well located for such work. Fig. 3

is an outline map showing the relative positions of these field receiving stations and the transmitting station.

The reason for settling down at a fixed point in this way was to attack the problem from a new angle. The field strength survey and aural observations had yielded much interesting information but did not appear at that time to shed a great deal of light on the quality distortion so it was decided to attempt, by an oscillographic

Fig. 2—Interior view of test station near Stamford, Conn.

study of received signals sent out under rigorously controlled conditions, to determine just what alterations these signals suffered in their journey through space.

In finding such distortions the ear is, of course, the primary testing instrument or indicator of trouble, for, if the trained ear is unable to detect anything wrong with a received signal in comparison with its original counterpart it is safe to say that nothing detrimental of importance has happened to it. But the ear is a poor quantitative indicator and furnishes no permanent or easily analyzed record of its observations. It is evident that if we are to study quantitatively the characteristics of radio transmission which give rise to quality distortion,

we must devise tests which will disclose changes, of whatever kind, in the relations between the various component frequencies of the transmitted band and furnish interpretable permanent records. In

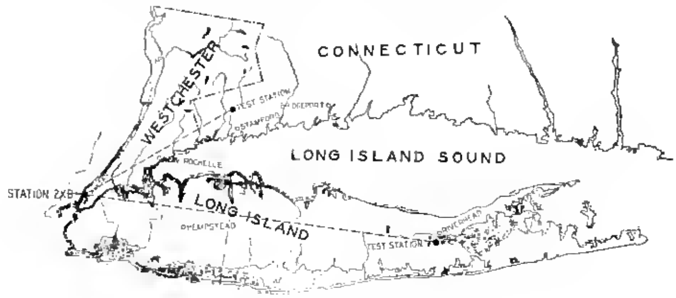

Fig. 3—Outline map showing locations of transmitting station and receiving test stations

fact in the studies described herein a considerable portion of the job was to devise or perfect suitable methods of attack.

Single, Double, and Triple Frequency Tests

The variable factors in radio transmission which may be directly controlled are located at the transmitter and receiver. We have as yet no tangible means of controlling the transmitting medium, but it can be studied indirectly through the characteristics of the received signals. Obviously, it is desirable in the interest of simplicity to stabilize the apparatus variables to the extent that they may be idealized in considering observed results. Furthermore, at both the transmitter and receiver, it is desirable to make the antenna arrangements of the simplest form. For our work the normal antenna arrangement at station 2XB was used perforce since any important changes would have constituted a major operation. It is far from a simple arrangement, as shown in Fig. 4 which is an outline elevation and plan of the antenna and building at 463 West Street, New York City. Fortunately there are no buildings considerably higher than the antenna within a distance of several wave lengths.

At the receiving test stations both loop and vertical antenna were used; but in most of the experiments a simple vertical antenna was employed. It was constructed of brass tubing, 30 feet long, and guyed in a vertical position. A galvanized iron pipe 12 feet long was driven in the earth for a ground connection. The vertical receiving antenna projected through the roof of the test station building

at Riverhead, L. I., as shown in Fig. 5. The receiving antenna was not tuned but was connected to the radio receiver through fixed inductive coupling.

The carrier power in the transmitting antenna normally remains fairly constant, except for minor variations in voltage of the supply mains, and with a little care on the part of operating personnel, the

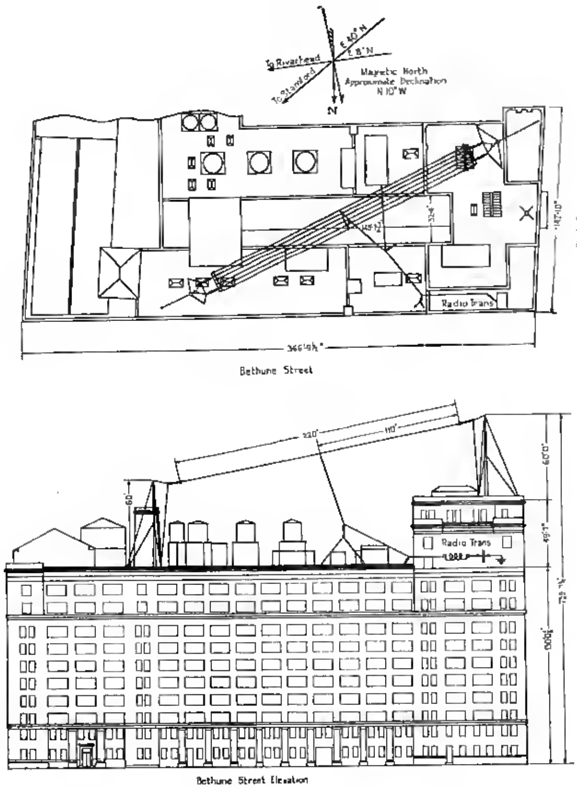

Fig. 4—Plan—elevation of the transmitting antenna

antenna current can be kept within the limits of a 1 per cent. variation, which is small compared with the signal fading usually experienced.

The stabilization of the frequency was of the greatest importance since in some of the tests it was desired to beat or heterodyne the signals down to audio frequencies and pass them through narrow band filters. To provide this stability engineers of the Bell Telephone Laboratories arranged the 5-kw. transmitter at station 2XB to obtain

its carrier frequency by amplification of the output of a 610-kc. piezo-electric crystal oscillator.

When desired some of the antenna current from the output of the transmitter was rectified and the resulting current was sent over a telephone line to the receiving station so that the frequency and wave

Fig. 5—Receiving test station near Riverhead, L. I. showing vertical antenna projecting through roof of building

form of the modulating signal could be seen and photographed at that point, thus guarding against any possible distortion in the transmitter and enabling a direct "before and after" comparison to be made. The telephone circuit was also used for communication between engineers at the two terminal stations.

At the receiving station double detection receivers and audio frequency amplifiers were employed. These did not have entirely "flat" transmission characteristics over the audio frequency band, but in most of the tests this was of no importance. In cases where it affected the results the making of necessary corrections was a simple matter. In tests involving beating the received signals down to audio frequencies through the agency of a local heterodyning frequency,

this was supplied from a shielded vacuum tube oscillator which on comparison with a standardized piezo-electric oscillator was found to possess the required stability. The double detection type receivers were used for no other reasons than their availability and their convenience for quantitative work. The beating down oscillator within the sets and the intermediate frequency step passed through in the sets by received signals do not figure in the following discussion of test methods but, of course, in each case the necessary set tuning

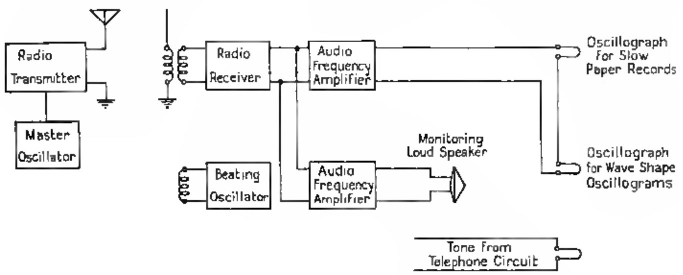

Fig. 6—Diagram of system used for single frequency tests

adjustments were made. To avoid confusion it is well to think of these receivers as being replaced by high frequency amplifiers and simple detectors since the local beating oscillator referred to in later pages is the separate shielded oscillator described above which is used to beat the signals down to audio frequencies.

In this work the moving coil type oscillograph was used throughout for the purpose of making photographic signal records. As indicated in Fig. 6 two oscillographs with elements connected in series were employed; one for the purpose of making a continuous record of the variation in the amplitude of the signal using a slow moving photographic paper tape and the other to obtain the wave shape of the signal by means of the usual high speed photographic film drum. An element of one oscillograph was also used at times to record on the film drum the wave shape of signals rectified at the transmitting antenna and sent over the telephone lines.

Fig. 7 is the interior view of the test station at Riverhead showing the general arrangement of the oscillographs and accessory apparatus. This oscillograph equipment formed about the only fixed portion of the apparatus, other portions being changed from time to time for different tests. These arrangements will be described later in connection with the records which they were used to obtain.

In considering these various records perhaps we had best look first at the simpler ones and then proceed in a more or less orderly

fashion to the more involved ones. The simplest records are fading records of the unmodulated carrier frequency of 610 kc. At the receiver the carrier was heterodyned with a local oscillator to produce a beat tone of about 250 cycles which was fed through amplifiers to the oscillograph elements.

A representative sample of the form of signal records made in the manner described above which show the variation in the amplitude of

Fig. 7—Interior view of Riverhead testing station showing recording apparatus

the received carrier signal with time, is given in Fig. 8. It shows a typical fading record made at Stamford, Conn., May 16, 1925. The timing interval on strip 6 is 2.6 seconds.

The feed of the photographic paper tape through the oscillograph was varied somewhat during the course of the experiments but was generally in the range of 6 to 12 inches a minute. At this rate the record of an audible frequency signal is a shadow band of varying

width corresponding to twice the amplitude of the signal, as both the positive and negative half-cycles are recorded. It will be observed that the outer limits of the band corresponding to the peaks of the sine wave are darker than the center portion of the record. This is due to the fact that the rate of change of the movement of the light

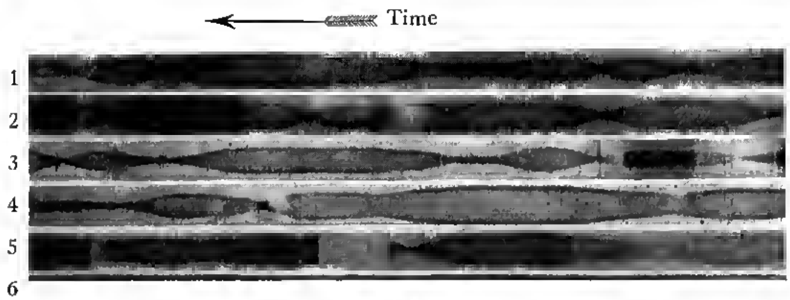

Fig. 8—Single-frequency fading record. Made at Stamford, Conn., May 16, 1924, 1:54 a.m. Timing marks, on strip 6, 2.5 seconds apart

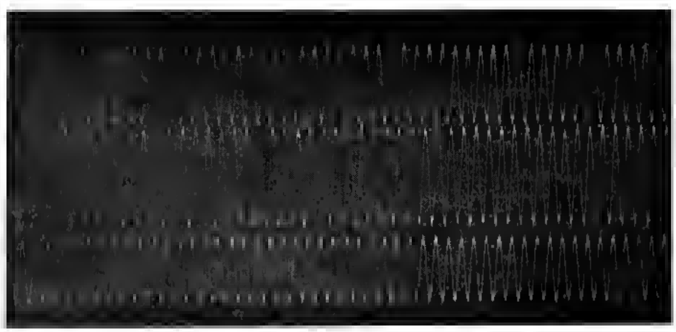

Fig. 9—Wave form of beat note signal for single-frequency test. Center trace signal from vertical antenna, upper and lower traces signals from loop antenna receivers

spot on the record is a minimum at the peak of the signal; hence, a greater quantity of light affects these portions of the record. This shading effect was very useful in the way it brought out changes in the distortion of the received signal. This is discussed fully in another section of the paper. The fuzzy irregular outline on portions of the records is caused by static and radio noise. The timing marks on the record allow a measurement of the time interval between points of minimum signal. Fig. 9 is a sample oscillogram of the wave shape of a beat note signal recorded by the method described above.

Marked changes in the fading cycle or time interval between points of minimum signal may occur within a period of a few minutes, and

from day to day there is often evidenced a modification of the general character and the recurrence of these changes. An example of this change in a short period of time is well illustrated by the oscillograms in Fig. 10. Strips 1, 2 and 3 form a continuous record starting at 1:52 a.m.; strips 4, 5 and 6 start at 2:16 a.m.; and strips 7, 8 and 9 start at 2:37 a.m. These are three sections of a continuous record

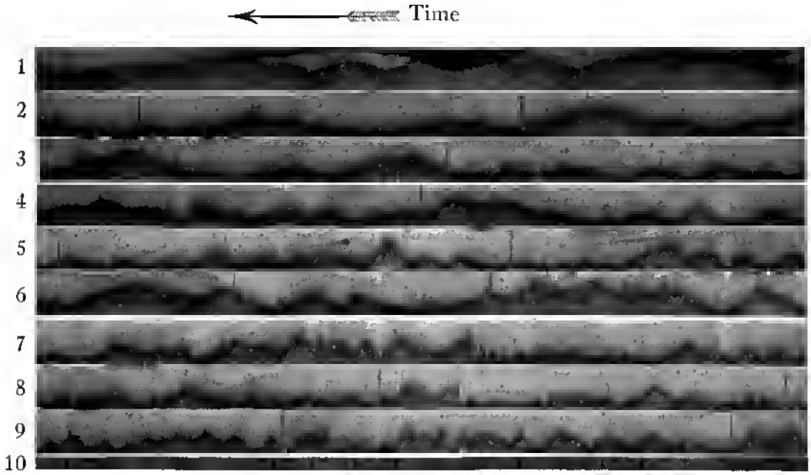

Fig. 10—Single-frequency fading record, showing variation in rapidity of fading, made at Riverhead, L. I., July 16, 1925, 1:52 a.m. Timing marks, on strip 10, 5 seconds apart

selected for the purpose of showing the decrease in the fading period, in a 45-minute interval. The timing interval on strip 10 which applies to these records is 5 seconds. In this particular record only half of the audio signal was recorded, the edge of the strip being the zero line.

These single frequency fading records do not offer very much to work on. There is, however, just enough suggestion of regularity about them to annoy one with the thought that perhaps they may follow some definite combination of periodicities and with this in mind we have taken sections of two different records and subjected them to a harmonic analysis.

So far we have been able to draw no more useful conclusions from such harmonic analyses than that the heterogeneous scattering of harmonic values is about what one would expect from the looks of the curves.

One significant thing about these oscillographic single frequency fading records is that they show no high speed fading of important

magnitudes. Occasionally one cycle of the beat tone will be somewhat upset by a sudden change in the amplitude but in general no changes which consistently distort the wave form were observed.

The slow fading may be considered as a modulation and on this basis the received signal is seen to be composed of the original constant carrier frequency accompanied by very narrow side bands occupying at best perhaps a fraction of a cycle.

The next progressive step in the radio transmission studies is naturally from a single frequency to two or more frequencies transmitted simultaneously. By the use of two crystal oscillators at the transmitter two separate and distinct radio frequency signals were transmitted simultaneously. These crystals were ground by the Bell Telephone Laboratories to oscillate at 610,000 cycles and 609,750 cycles. The amplitudes of these signals at the transmitter were controllable so that it was possible to make them equal, or one larger than the other, equivalent to the relative magnitudes usually found for the carrier and single side-band transmission case. Records were obtained of the variation of these radio signals, but none is reproduced here since the information shown by them can be just as easily obtained from the triple frequency records shown below.

Radio transmission on three frequencies is readily obtained by modulating the carrier with an audio frequency tone, and observing the three frequencies separately at the receiver.

If the modulating tone is

$$\sin (vt+\phi)$$

and the carrier signal

$$A \sin pt,$$

the transmitted signals are

$$+\frac{Aa}{2} \cos [(p+v)t+\phi] \qquad \text{(upper side band)}$$

$$+A \sin pt \qquad \text{(carrier)}$$

and

$$-\frac{Aa}{2} \cos [(p-v)t-\phi] \qquad \text{(lower side band).}$$

where a is a constant proportional to the percentage modulation.

These three frequencies are not merely a mathematical fiction but are physically existent as three separate waves bound together only at their point of origin.

To adequately record them separately by means of the oscillograph advantage was taken of the fact that a group of frequencies beaten

with a single frequency differing from them by a small amount and detected may thereby be reduced to audible frequencies without having their interrelations of phase, amplitude or difference frequency composition, changed in any respect. For instance if the frequencies expressed above are beaten with a local constant frequency,

$$B \cos (qt+\psi)$$

the resultant lower or difference frequencies will be

$$\begin{aligned} &+\frac{kBAa}{2} \cos [(p+v-q)t+\phi-\psi] \\ &+kBA \quad \sin [(p-q)t-\psi] \\ &-\frac{kBAa}{2} \cos [(p-v-q)t-\phi-\psi]. \end{aligned}$$

Each one of the three components has been changed in amplitude by the same factor kB representing the efficiency of detection. Each

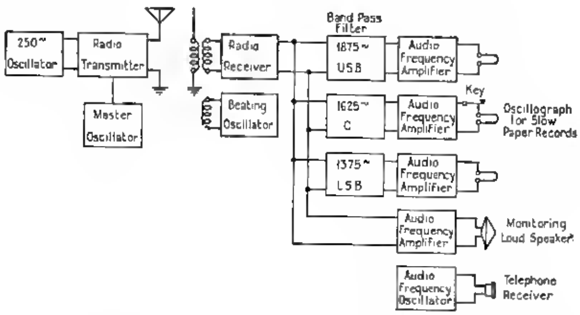

Fig. 11—Diagram of system used for three-frequency tests

one has been reduced in frequency by exactly the same amount $\frac{q}{2\pi}$ and each has had its instantaneous phase shifted by an angle $-\psi$. Relative to each other they remain unchanged.

In our actual case the carrier frequency $\frac{p}{2\pi}$ was 610 kc. The modulating frequency $\frac{v}{2\pi}$ was 250 cycles and the beating frequency $\frac{q}{2\pi}$ was 608,375 kc. so that the resulting three audio frequencies were 1,875 cycles, 1,625 cycles and 1,375 cycles.

As indicated in Fig. 11 in order to make a record of these signals they are separated at the receiver by means of band filters. These filters and others similar in type for other modulating frequencies

were designed and made by the Bell Telephone Laboratories especially for this work. The band filters used for the purpose of selecting the carrier and side-band frequencies had a cutoff of 40 Transmission Units 250 cycles from the mid-band frequency.

These cutoffs together with the position in the frequency range of the pass bands of the filters preclude any troubles from cross modulation of the radio carrier and side bands during the beating down process. The products of such cross modulation would be frequencies which are multiples of 250 cycles and these cannot pass the filters. On the other hand the beaten down frequencies will pass practically intact, since as has been shown by the previously described single frequency tests, each of the three frequencies received although subjected to amplitude modulation by fading, represents only a very narrow band of frequencies for which the filter pass bands were of adequate width.

As the modulating tone was carefully calibrated to 250 cycles and the filters adjusted to transmit the frequencies specified, it was only necessary to transmit the carrier while adjusting the receiving beating oscillator. The following procedure for this adjustment was found to be very successful. A local audio frequency oscillator was set to the reduced carrier frequency of 1,625 cycles, and its output connected to a telephone receiver. The audio beat note from the radio signal and local beating oscillator was reproduced by a loud speaker and its frequency adjusted to zero beat the 1,625-cycle tone from the telephone receiver.

When this adjustment had been completed the carrier was modulated with the 250-cycle tone, and the side-band signals automatically pass through their respective filters.

The signals from the outputs of the filters were amplified, and recorded separately by the three oscillograph elements. The sample records shown in Fig. 12 are representative.

Strips 1, 2 and 3 are taken from a long record obtained May 7, 1925, 3:22 a.m. The upper trace is a record of the upper side-band signal, the center trace the carrier, and the lower trace the lower side-band. Strips 4, 5 and 6 are from a section of a similar type of record made May 23, 1925, 1:06 a.m., where the carrier was modulated with a 500-cycle tone and different filters were used. In this record the upper trace is the lower side-band and the lower trace the upper side-band.

It will be noticed that the timing interruption appears only in the side-band signals, as the tone was interrupted before modulation took place, and that the amplitude of the carrier signal is not affected

by the interruption of the modulating tone. This makes it very easy to identify the side-band signals. These records give an excellent graphic picture of ordinary radio telephone transmission, bringing out the fact that three truly individual frequencies are transmitted to reproduce one.

In Fig. 12, strips 1, 2, and 3, the relative amplitudes of the three signals are very nearly in proportion to the relative amplitudes of

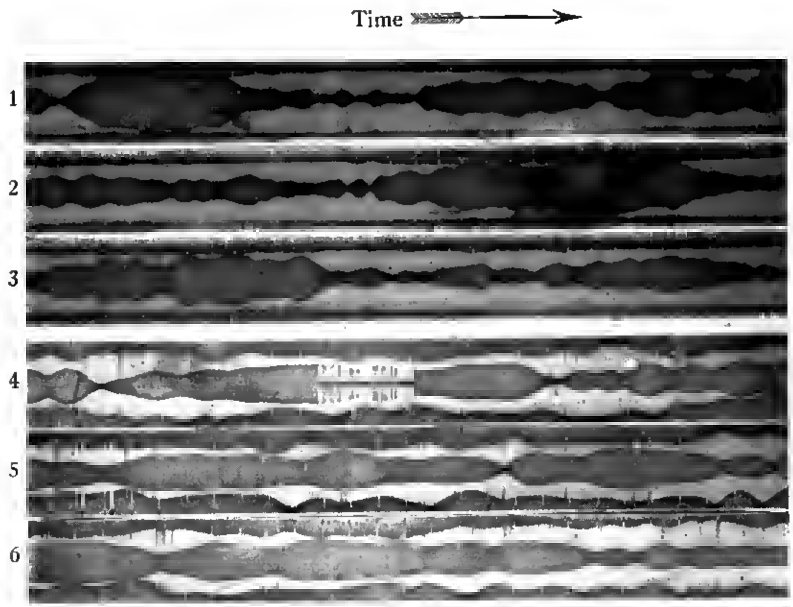

Fig. 12—Fading record showing individually the fading of carrier and side-band frequencies. Made at Riverhead, L. I. Timing interruptions in side-band signals, 5 seconds apart

the signals as they existed in the ether at the receiving point. Before this record was made a transmission characteristic of the complete receiving circuit, including the oscillograph elements, was obtained, using a local transmitter with modulated carrier for the purpose of making the measurement. The gain of the audio amplifiers at the outputs of the filters was adjusted to give substantially uniform transmission on each of the three frequencies corresponding to the carrier and side bands of the radio frequency signal.

As shown in Fig. 11, a telegraph key is placed in the circuit of the center oscillograph element, for the purpose of placing identifying signals on the records. An example of these identifying signals is

shown in Fig. 12, strip 4, which gives the date and time the record was started, July 23, 1925, 2:06 a.m. (Eastern daylight saving time).

The record in Fig. 13 is of the carrier and side-band signals with 500-cycle modulation made at Riverhead, L. I., May 25, 1925, 1:25 a.m. More gain was used in the side-band amplifiers for this record in order that the effects of fading could be brought out more prominently. In this record only half of the side-band signals were recorded, the

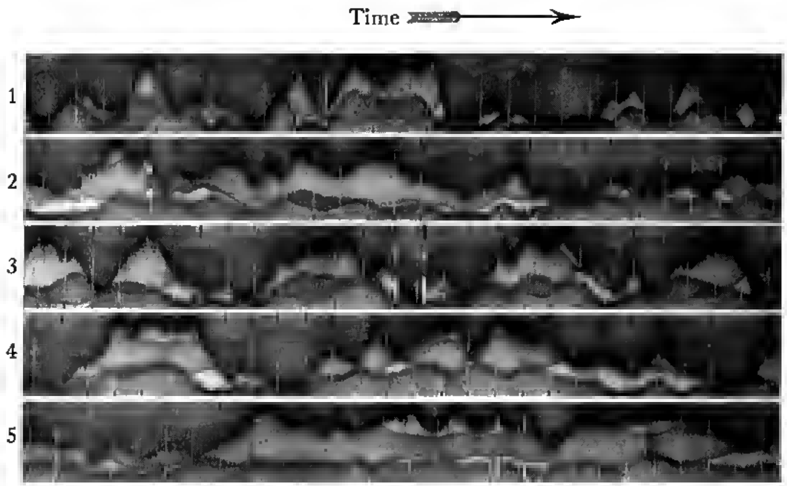

Fig. 13—Fading record of carrier and side-band signals, made at Riverhead, L. I. Timing interruptions in side-band signals, 5 seconds apart

zero reference line being at the edge of the strip. The upper trace is the upper side band, the center the carrier and lower trace the lower side band. Where the traces of the signals overlap a darker record is obtained. This record may be confusing at first but if strip 5 is examined where the amplitudes of the signals are not so large a better picture of the form of the record will be obtained.

It is obvious from these records that the carrier and side-band signals do not fade together as a unit. The carrier may pass through a zero value with still considerable amplitude in the side-band signals as in strips 1 and 3. In the first case, strip 1, the three frequencies successively fade through points of minimum signal in the order lower side-band, carrier and upper side-band; and in the second case, strip 3, the three frequencies fade through points of minimum signal in the reverse order. This is a definite indication of *selective fading*; that is, *fading is a function of frequency as well as time.*

An endeavor to form an explanation of the cause of this selective action in fading must be largely in the nature of speculation. Furthermore, since our data consist in the results of things which have happened rather than in any first hand information on the processes of the happening, the building of an explanation is a synthetic process. In general for any given set of facts it is possible to synthesize a number of explanations. Bearing this philosophy in mind we have considered various theories in connection with our observations and have concluded that simple wave interference as a major cause of the signal variations is at present the most likely explanation. While wave interference may be called a major cause it should perhaps also be called a secondary cause since the assumption of wave interference presupposes for its origin, primary causation by some physical state or configuration of the transmission medium. Speculation as to the nature of this primary cause is one stage further removed from the data contained in our oscillographic records than is the assumption of wave interference.

Since it is desirable in the remainder of this discussion to point out the evidences of wave interference, let us consider briefly the nature of this phenomenon.

To avoid any possible confusion of terms let it be said that by "wave interference" we mean a particular physical phenomenon in wave transmission and have no reference whatever to static, signals from other stations, or any other of the forms of radio noise which are commonly designated by the word "interference" when they hinder the reception of desired signals.

When two single frequency plane polarized wave trains start out at the same time from a common source and travel by different routes to meet again at a distant point the nature of disturbance at that point is determined by the relative space phases of the planes of polarization and time phases of the amplitude of the two arriving waves.

If we let E represent the vertical resultant of the electric field, which would be the only part affecting a simple vertical antenna, such as we have used in most of our tests, then

$$E = e_1 \sin 2\pi(Ft+d_1) + e_2 \sin 2\pi(Ft+d_2) \qquad (1)$$

where F is the frequency and d_1 and d_2 are the distances along the respective paths measured in wave lengths and e_1 and e_2 are the vertical components of the two waves. These two sine terms may be thought of as two vectors differing in phase.

The condition that these add giving a field

$$E=(e_1+e_2)\sin 2\pi Ft$$

$$\text{is that, } d_1-d_2=(\text{a whole number}) \tag{2}$$

that is, the difference in length of the two paths must be an exact whole number of wave lengths. The condition that the two waves cancel each other giving a field

$$E=(e_1-e_2)\sin 2\pi Ft$$

$$\text{is that, } d_1-d_2=(\text{a whole number})+\tfrac{1}{2} \tag{3}$$

that is, the difference in length of path must be an exact odd number of half wave lengths.

Thus if the two components e_1 and e_2 are equal, the resultant vertical field E will go through values ranging from (e_1+e_2) down to zero as the path lengths change relative to each other. If the two waves do not have exactly the same amplitude, the minimum value of E will be something more than zero.

Differences in attenuation of the two waves and differences in their direction of arrival will modify the relative amplitudes of e_1 and e_2 but will not modify the time relations required for minima of the resultant field E unless we assume that at the time of a minimum neither wave has an appreciable vertical component. Since the consequences of such an assumption do not accord with our experimental data we have considered that it may be left out of account in the present discussion.

This is obviously a picture which fits in very well with the simple single frequency fading records. The major maxima and minima occur when the conditions of equations (2) and (3) are met and e_1 and e_2 are nearly equal. On the other hand it seems doubtful that the picture can be so simple. If we suppose two wave paths why not three or more? Additional paths would add irregularities to the fading and it would not be necessary to assume as great a degree of irregularity in the changes in any one path. But with an increasing number of paths the various arriving waves would tend to average to a more or less constant mean value and large departures from this mean would become rare. The fact that the fading signal continually covers a large range of amplitude, with the maximum many times the minimum, definitely points toward there being but a very small number of major paths, probably not more than two.

Considering now the question of selective fading in relation to wave interference we refer back to equation (2).

If we assume the distances to be measured in any desired units and call them d_1' and d_2' our equation will still hold provided we divide each distance by the wave length measured in the same units, thus

$$\frac{d_1'-d_2'}{\lambda}=\text{a whole number}=x;$$

rearranging this and writing $\frac{V}{F}$ for λ where V equals the velocity of the waves, we have

$$F=x\left(\frac{V}{d_1'-d_2'}\right). \tag{4}$$

If now we assume $(d_1'-d_2')$ to be fixed we find that F can have a series of values which are integral multiples of $\frac{V}{d_1'-d_2'}$ which we may call the frequency spacing interval. That is, with changing frequency E will go through maximum values with frequency at a series of frequencies beginning theoretically with zero and extending upward in regular spacing to infinity.

The spacing interval is obviously that number of cycles which corresponds to the lowest finite frequency in the series, namely, the frequency for which the distance $(d_1'-d_2')$ is just one wave length since when $x=$unity equation (4) becomes

$$F_1=\frac{V}{d_1'-d_2'}=\text{the spacing interval.} \tag{5}$$

By using the same process on equation (3) we find that E has minimum or zero values at another series of frequencies having the same spacing interval but lying midway between the frequencies at which maxima occur.

Thus it is apparent that with fixed path length difference the amplitude of the field E will be different for different frequencies, ranging from maxima of (e_1+e_2) down to minima of zero if the polarization planes and amplitudes of the two vertical components are equal.

Furthermore, still thinking of equation (1) as representing two vectors, it is evident that the phase of the resultant field is different for different frequencies even though these different frequencies had exactly the same starting phase at the source.

If the paths are changing with time, the field at a given point, as has already been pointed out, will go through time fluctuations. Another way to look at this is that there is a space pattern of maxima

and minima and as the paths change the plane section of the pattern taken by the surface of the earth wanders so that at any one point the field is continually fading in and out as the maxima and minima glide by it. Each frequency has its own pattern differing from those of its neighboring frequencies in such a way that at any given point the relation between amplitude and frequency is that just discussed

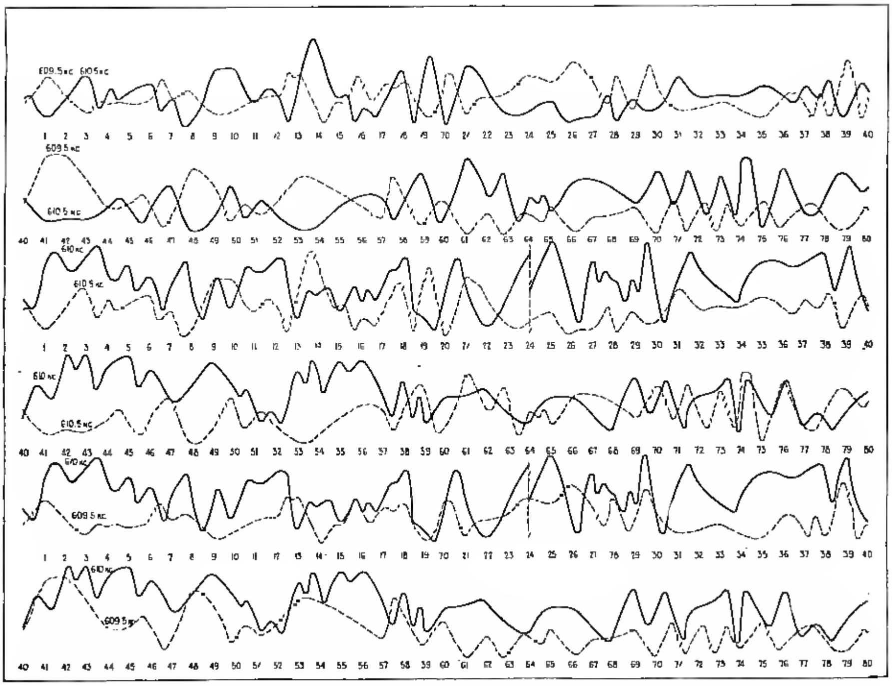

Fig. 14—Plotted curves of signal amplitudes condensing a long fading record, part of which is shown in Fig. 13. Numbers along time axis correspond to successive 25 second timing interruptions

above. Thus as the paths change and the patterns shift the different frequencies fade not simultaneously but progressively.

In the above analysis of wave interference it has been assumed that all frequencies traveled from transmitter to receiver over a given path in the same elapsed time. This does not mean that they necessarily follow exactly the same route on this path since they might follow somewhat different routes of equal length or if their transmission velocities were different they might follow different routes of unequal length and still come within the definition of a "path." It seems reasonable to assume that over the width of an

ordinary transmitted band the various frequencies are treated alike by the medium and the simple assumption that they follow the same route with the same velocity is justified. If none of these assumptions is correct but the departure is not large the effect will be merely to introduce slight irregularities into the spacing interval and the general nature of the result will not be changed.

Let us now examine more closely the record, a part of which is shown in Fig. 13. A portion of this has been condensed into the curves of Fig. 14. One unit along the time axes of these curves represents a 25-second interval.

To obtain these curves the amplitude of the signal has been scaled off and plotted, ignoring all the minor irregularities. From this record the relative fading characteristics of these single frequency signals 500 cycles apart are more easily seen, and it is possible to contrast the time of occurrence of points of minimum signal for any pair of them.

For the frequency difference of 500 cycles (610.5–610 and 610–609.5) these times are obviously quite different but there is no clearly

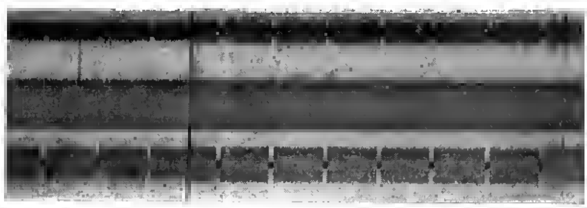

Fig. 15—Daytime record of carrier and side-band signals

discernible relation between them. The curves for 1000-cycle difference (609.5–610.5), however, show a striking relation in that the maxima and minima of the two are opposed fairly regularly over the entire 33-minute interval covered by the plot. This means that when one frequency has a minimum amplitude the other has a maximum and vice versa. Certainly this suggests a wave interference involving only two major paths whose difference in length is such that the spacing interval is 2,000 cycles. The path difference appears to be changing somewhat irregularly but at an average rate of the order of one wave length (or approximately 500 meters) per minute.

Before speculating further on the numerical values which may be derived from this part of the data we had perhaps best consider some other records of a somewhat different kind which are better adapted to provide such values. But first let us reiterate that these are *night-time* effects.

During the day signals substantially uniform in amplitude are received. An example of the type of transmission obtained in the daytime is given in Fig. 15, which is a record of the carrier and side-band signals received with substantially the same terminal conditions with the exception of the time as that existed when the records shown in Fig. 12 were made.

The abrupt change in the amplitude of the side-band signals was due to an intentional change at the transmitter in the input level of the tone modulating the carrier, and accordingly the amplitude of the carrier did not change. The timing interval is 5 seconds.

Band Fading Records

The familiar fading record is limited to two axes, amplitude and time. So far we have extended this cramped perspective somewhat by observing as many as three separate fading records spaced at audio-frequency intervals along the frequency axis. Even these three narrow lookouts upon the wide range of ether transmission have indicated amplitude relations along the frequency axis which

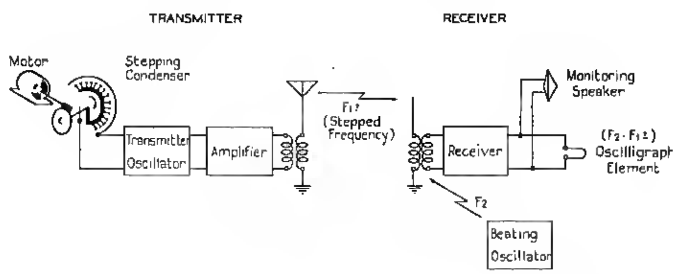

Fig. 16—Diagram of system used to obtain records of selective fading or "band fading" records

promise to open a new line of attack upon the problem of night-time fading. But the desirability of knowing what takes place in the interval unrevealed by these cracks in the fence becomes obvious. We should like to know the relative amplitude of frequencies over a wide band, and the change in this relation with time.

Since it is not a simple matter to record simultaneously the amplitude of a large number of waves of frequencies separated by say one hundred cycles in the radio-frequency range a single frequency in combination with a frequency stepping device at the transmitter has been adopted. The circuit arrangement is shown diagrammatically in Fig. 16. The rotary contactor bringing into the circuit suc-

cessively a total of fifteen small condensers across the main condenser of the transmitter oscillator shifts the frequency in steps over an adjustable range. The contactor is rotated at the rate of nine revolutions a minute, which is sufficiently slow to show definite steps in the oscillograph record. At the receiving end a local oscillator supplies a radio-frequency wave for beating the incoming frequencies down to values within the audible range.

A long oscillograph record of this stepped frequency gives a sort of moving picture of the fading for the entire band covered. A sample of such a record is shown in Fig. 17 with alternate pictures in the series removed to simplify the relations, since by reason of

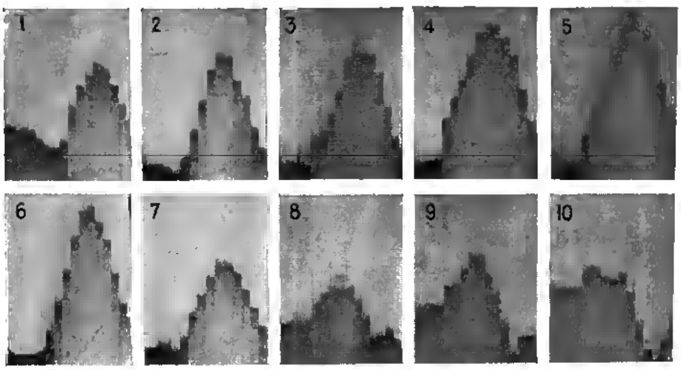

Fig. 17—Sample band fading record

the two-way traversal of the frequency band successive pictures are reversed. If a series of such built-up pictures as these could be taken rapidly on moving picture film, and projected successively upon a screen we should have before us an animated view of band fading. And according to the results of experimental investigation the subject offers a lively theme for such a presentation. The peaks and depressions glide nervously back and forth across the setting. The successive pictures of Fig. 17 (which, by the way, were selected for their half-tone reproduction possibilities rather than as first class examples of the records taken) illustrate a rather leisurely movement of this sort. These ten built-up photographs cover a period of slightly more than one minute. In the first seven pictures a depression appears at the left, while in the last three this depression seems to have made an exit followed by the simultaneous entrance of another from the opposite wing of the stage. Evidence of such

organized spacing of the minima is present in all of these night-time band fading records. As has already been suggested such evidence has an important significance, but before going into this phase of the subject again let us examine a little more in detail the structure of these band fading records.

The steps in any one picture of Fig. 17 are, as we have said, snap-shots of the wave amplitude for successively different radio frequencies taken about a quarter of a second apart. The fact that the fifteen snap-shots used to build up a single picture are not taken simultaneously causes a skewing of the outlines when movement of the depressions as shown in Fig. 17 occurs. If, for example, we

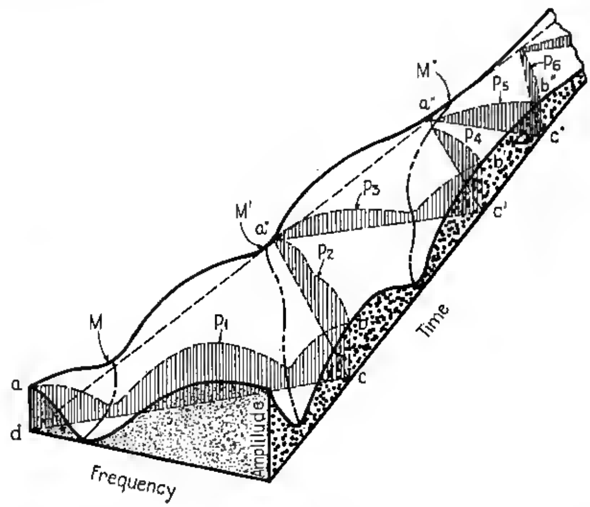

Fig. 18—Three dimensional diagram, showing the method of interpreting band fading records

were to take fifteen separate and successive snap-shots of a mountain through fifteen long vertical slits side by side it would be possible to combine the narrow sections so as to form a true picture of the peak. Now, if by some prodigious act of nature the mountain were shifted suddenly to one side and back again during the time we were taking the fifteen successive snap-shots through the vertical slits, the combination of them would form a profile quite different from that obtained when it was stationary. Or if it were simply moved steadily across the field of vision during the time the snap-shots were being taken one slope would be made to appear precipitous while the other would be leveled to a gentle grade in the finally built-up picture.

The character of this skewing, then, and its magnitude depend upon the rate at which the object being photographed in vertical sections moves, and the direction of the movement.

In Fig. 18 is shown an imaginary night-time band fading record in the "assembled" form. Since such a record contains frequency as a third dimension, in addition to amplitude and time as shown in the ordinary fading record, our simple fading curve has assumed the broader aspect of a surface, the selective fading making more or less parallel "valleys" running across it. The step-frequency system of recording the points amounts to photographing sections of this solid. The important point to be kept in mind is that these sections are *not perpendicular to the time axis.* If they were, the skewing previously described would not be present. By setting these sections up in their true relation to the time axis, however, and filling in to produce a continuous surface such as is shown in Fig. 18 the result is correctly represented. In order to make a detailed and accurate study of the band fading records, therefore, it is desirable to construct from the oscillograph sections the complete surface by the method suggested.

In Fig. 18 the trace of minima crossing the band is shown by M, M' and M''. Picture sections obtained as our recording apparatus literally moves back and forth across this frequency band are shown as $(a\text{–}b\text{–}c\text{–}d)$, $(b\text{–}c\text{–}a')$, $(a'\text{–}b'\text{–}c')$, etc. It will be evident that the section P_1, for example, will, in case a minimum is crossing rapidly, appear entirely unrelated to section P_2. When the minima run nearly parallel to the time axis (slow changes in transmission conditions) the successive pictures P_1, P_2, P_3, etc., will reveal their relation by direct comparison.

Actually to obtain frequency-amplitude sections perpendicular to the time axis in Fig. 18 would require the simultaneous transmission and reception of a large number of frequencies spaced at short intervals along the frequency axis. A more practical thought is to speed up the process and though this seems very simple at first consideration, it will be shown later to involve a particular kind of distortion which cannot be separated out as easily as the skewing encountered by the more deliberated method.

Now that we are familiar with the data, Fig. 19 showing, partially superimposed in vertical strips, the outlines of successive built-up pictures of the frequency traverse will be of greater significance. During the steady periods there appears within the 2,280-cycle band covered by these data approximately one complete cycle of selective fading. The lack of flatness in the audio-frequency-transmission characteristic of terminal apparatus has caused the suppression

of amplitudes toward the right side of these sections. Keeping in mind also the skewing inherent to this system of presentation during transient periods, we are able to trace the movement of minima, as illustrated previously in Fig. 17 which was taken from a different

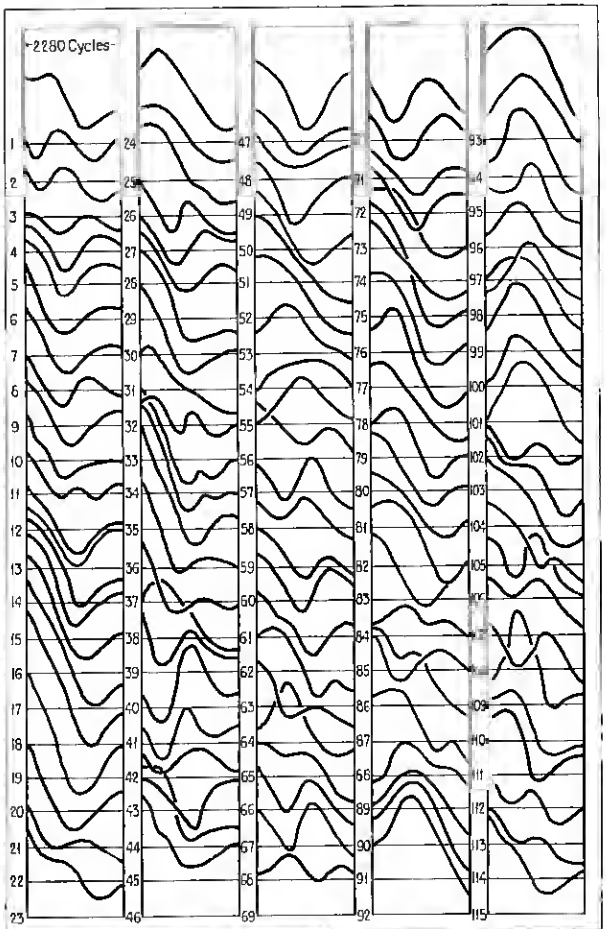

Fig. 19—Plotted curves, condensing a long band fading record so as to bring out the effect of selective fading

record. The relative position of these minima gives us an interesting insight into the nature of the night-time transmission path.

From records covering frequency ranges up to 4,500 cycles in width the positions of major minima along the frequency axis have

been plotted against time as in Fig. 20. The widths of the frequency bands covered in this case are indicated. This picture is essentially a bird's-eye view of band fading records such as are illustrated in idealized form by Fig. 18, the amplitude axis being perpendicular to the page. It reveals the presence of minima spaced at more or less definite frequency intervals, and suggests the presence of other

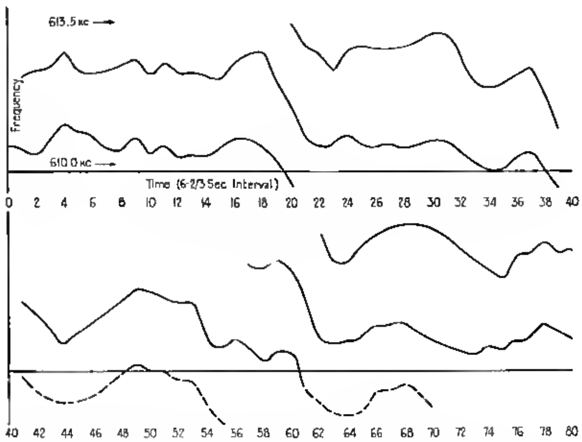

Fig. 20—Plotted curves which condense a long band fading record so as to bring out the frequency spacing interval of the selective fading

depressions in regular spacing beyond the scope of our pictures, for when one minimum slides out of sight another appears to take its place from the opposite side of the band. The minima traces shown in broken line were outside the record but were located by extrapolating the sections.

Other depressions of small amplitude appear to be superimposed upon the major changes but the present data appear inadequate to give reliable information concerning them. These minor depressions seem most evident during periods of rapid change.

The presence of these major minima in regular array bears a marked similarity to the familiar wave interference case in light and fits in very nicely with the theory detailed in previous paragraphs. Assume for a moment the simple case of two transmission paths producing such an effect and account for the difference in their lengths by presuming that one path follows more or less closely along the surface of the earth while the other seeks higher altitudes and in some fashion gets back to earth at the receiving station.

The mean frequency difference or spacing interval between successive minima for the records given in Fig. 20 is approximately 2,200 cycles. Therefore, the mean wave length difference in length of path from equation (5) is 277 wave lengths, or 136.5 kilometers.

It is evident that the errant waves following the second path must have been led a devious route. While this is about all the information which can be deduced directly from these data it is interesting to speculate further with the information along the lines of some of the theories which have been proposed to account for such wave deflections. For instance there is the Heaviside layer theory in which there is supposed to be a more or less well defined reflecting layer in the upper atmosphere. For this we would visualize our high altitude waves as proceeding in a straight line up to the layer, being reflected, and striking back to earth at the receiving station.

Since the distance from transmitter to receiver was 110 kilometers the length of the secondary path was 110+136.5 or 246.5 kilometers. By triangulation the height of the assumed reflecting layer may be determined as very nearly 110 kilometers or equal to the distance from transmitter to receiver, and the angle of incidence is 26.5 degrees.

As yet no positive information has been acquired concerning the variation of difference in length of two major night-time transmission paths with direct distance from the transmitter. If the path difference is due to reflection from an overhead layer, the expected relation by triangulation becomes quite simple.

$$\Delta d = \sqrt{\frac{y^2}{4}+h^2}-y.$$

When Δd is the difference in length of path, y is the direct distance and h is the vertical height of the layer.

An investigation of this relation would probably do much to prove or disprove the reflection theory.

At this point it is well to recall the results of earlier tests in which it was observed that single frequency waves separated by 1,000 cycles faded in approximately an inverse relation also indicating a spacing interval of about 2,000 cycles. The agreement of these earlier records is particularly noteworthy since about three weeks elapsed before the more detailed band fading records were made.

Fig. 20 shows a time variation in the frequency position of the minima which is explained as due to a variation in the difference of path length. If we indulge in further speculation along the line of layer phenomena we conclude that the reflecting layer is rising and falling. It is improbable that the whole layer would rise and fall

together so we conclude that undulations occur along its surface. These undulations in themselves would cause the length of path of the wave reflected toward the receiver to undergo a continual change. They would also introduce minor reflections from surfaces more distant than that responsible for the major effect which may be responsible for the more rapid, low amplitude fading which is usually superimposed upon the slow changes. Obviously, the character of the fading would in the event that it is caused by undulations along the reflecting layer, be determined by the amplitude and direction of movement over the surface.

If, on the other hand, we examine the possibilities of theories such as those proposed by Nichols and Schelleng, Larmor and others in which the action of free electrons in the atmosphere is invoked we might visualize the waves on the second path as following a curved trajectory. Or we might have the two sets of waves start off together, become split by double refraction and eventually come together again. Perhaps their planes of polarization will have been rotated. In fact it is possible to build up what appears, we must confess, to be a highly imaginary explanation in which the wave interference is accounted for not on the basis of any great difference in path length but by the assumption that the amount of rotation is such a function of frequency that a change of about 2,000 cycles adds or subtracts a complete rotation, and the further assumption that one set of waves has had its plane of polarization rotated through several more complete rotations than has the other. The synthetic possibilities are almost endless and we must wait upon further data more varied in character before the facts can be established. In the present investigation we have not attempted to determine the mechanism of the transmission medium except insofar as it could be inferred from the results of our tests which were aimed at finding out just how radio signals look after they have been subjected to a trip through this mechanism.

Returning to the solid band fading record illustrated in Fig. 18, let us form some conception of the appearance of this figure were it extended toward the much higher and lower frequencies using as a basis of this conception the supposition that the existing record is systematically distorted by wave interference. For a given rate of change in the physical difference in length of path, such as would be encountered in the simple reflection case, the rate of movement of the minima across the band fading pictures would vary directly with the frequency. Therefore, we can extend the narrow section shown in Fig. 18 to form a wide band fading record such as is shown in

Fig. 21, wherein we are looking down upon the distorted surface, the minima being traced by the light lines. Toward the short wave end of the band it is evident that a fading record for a single frequency represented, for example, by a section parallel to the time axis and perpendicular to the page, $a-a'$, would show rapid fading, while a similar record at the long wave end of the range as $b-b'$ would give slow amplitude changes. Such sections representing theoretical

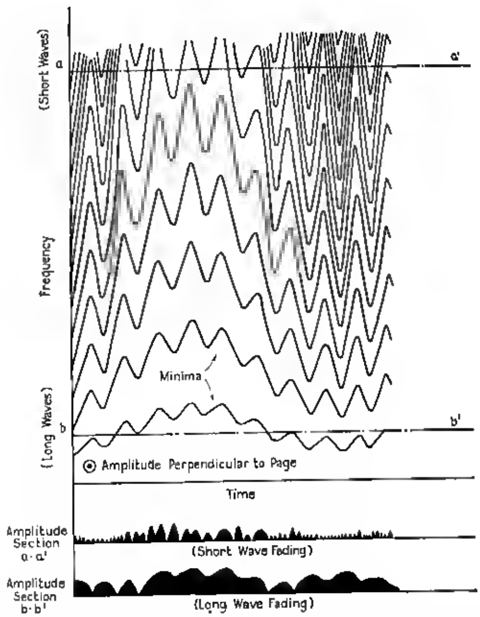

Fig. 21—Theoretical diagram obtained by extrapolating band fading records to show how the rapidity of fading might be expected to change with the wave-length

single frequency fading records are shown at the bottom of Fig. 21. The relative fading rates for long and short wave lengths as indicated by these idealized characteristics, are in accord with general experience

In describing the stepped-frequency method of obtaining band fading records allusion was made to distortion which might result from speeding up the process. Suppose that we were to use a very small rotating condenser in parallel with the main condenser of the transmitter oscillator for changing the frequency, and that this condenser were capable of changing the frequency sinusoidally about a mean value. Then we could represent the variation in frequency with time as is shown by the curve C_1 in (a) of Fig. 22. Now if the energy

transfer from transmitter to receiver takes place over two paths of different lengths one wave will constantly lag behind the other.

This lag may be measured as a time interval. In Fig. 23 are shown two waves, (a) and (b) of constant amplitude but with frequency

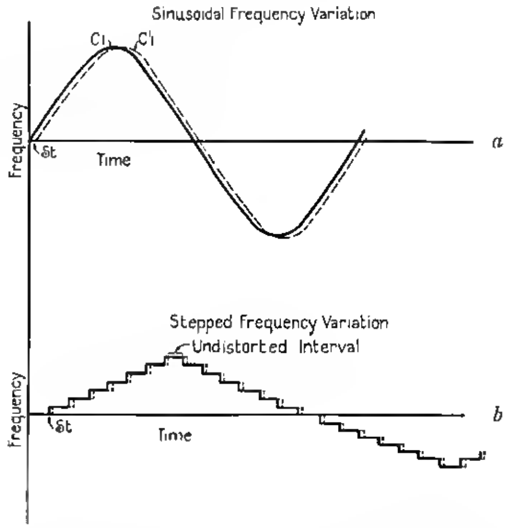

Fig. 22—Curves showing the relative effect of transmission time lag in sinusoidal and step-by-step methods of frequency variations

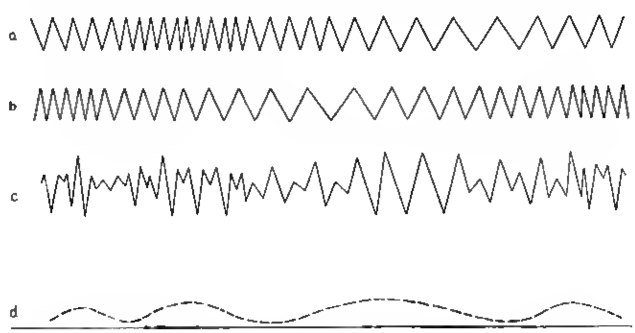

Fig. 23—Diagram showing the effect of frequency modulation

modulation. The wave (b) representing the indirect wave, it will be noticed, lags behind the direct wave represented by (a). The amount of this lag is determined by the difference in length of path and the transmission velocity. If we were to receive only one wave, as we should in the daytime, for example, we would find it to be a

constant amplitude field (providing the high-frequency characteristic of the receiver is flat over the range of frequency variation). But when two or more distinct paths exist, the combination at the receiver becomes complex. This is evident in curve (c) shown in Fig. 23 which is a direct summation of (a) and (b), and in (d) which is the envelope of (c). The amplitude is subjected to variations which did not exist at all in the original wave.

We might set up an *equivalent* effect right at the receiver by constructing two small local oscillators having the same characteristics as the transmitter oscillator. The two small rotating condensers would be driven by the same motor but the rotor of one would be shifted backward in phase relation to the other so as to simulate the case of transmission lag over the longer path. The relative frequency characteristics of the two may then be represented by curves C_1 and C_1' in (a) of Fig. 22.

The frequency of the signals arriving over devious paths at the receiver may be put in the form of an equation as,

$$F_1 = F_o + f \sin [r\ (t - d_1/V)], \tag{6a}$$

$$F_2 = F_o + f \sin [r\ (t - d_2/V)], \tag{6b}$$

wherein,

F_o = the mean frequency

f = one-half the total variation

$r = 2\pi$ times the frequency of rotation of the condenser

d = length of path

V = velocity of waves.

For a difference in length of path equal to 300 wave lengths at a frequency of 600,000 cycles per second, for example, the time lag of one wave behind the other will be equal to 300/600,000 second or 1/2000 second. The lag of one of the condensers behind the other in the "equivalent" case described above would be then for 30 cycles per second rotation of the condensers, 30/2000 times 360 degrees or 5.4 degrees. The lag of 5.4 degrees represents the lag of the condenser rotor so that the frequency lag will depend entirely upon the rate of change of frequency by the rotating condensers at any given instant.

Now to determine the resultant wave at the receiver we must know both amplitude and relative phase of the components arriving over the different paths. The amplitude will be constant, and we shall assume known, although it may actually follow slow changes with

attenuation or variations in length of path. The relative phase must be determined from equations (6a) and (6b). Knowing the frequency variation with time we may by integrating the following equation determine the phase relation at any time (t).

$$\Theta_1=\int_0^t 2\pi F_1 dt, \tag{7}$$

$$\Theta_2=\int_0^t 2\pi F_2 dt. \tag{8}$$

Substituting the general relation for F_1 and F_2 from equations (6a) and (6b) we have,

$$\Theta_1=\int_0^t F_o+f\sin r\ (t-d/V), \tag{9}$$

$$\Theta_2=\int_0^t F_o+f\sin r\ (t-d'/V). \tag{10}$$

Evidently the relative phase ($\Delta\Theta$) will be the difference between these two giving,

$$\Delta\Theta=\Theta_1-\Theta_2=2\pi\int_0^t F_o\,dt+2\pi\int_0^t f\sin r\ (t-d/V)\,dt \tag{11}$$

$$-2\pi\int_0^t F_o\,dt-2\pi\int_0^t f\sin r\ (t-d'/V) \tag{12}$$

which integrated reduces to the form,

$$\Delta\Theta=\frac{2\pi f}{p}\ (\cos r\ t-1)\ (\cos r\ d'/V-\cos r\ d/V+$$
$$\sin rt\ (\sin rd'/V-\sin rd/V). \tag{13}$$

The equation is not in itself very illuminating, but what it tells us generally is that if we represent two frequency modulated waves travelling over paths of different lengths to a distant receiver by rotating vectors, these vectors are constantly shifting their relative position. The magnitude of the shift at any instant is given by the varying angle $\Delta\Theta$. Due to a change in the angle included by the two vectors their resultant will undergo an amplitude change, the seriousness of which we will consider later.

Thus far in the discussion of frequency modulation by means of a rotating condenser we have assumed sinusoidal changes in frequency. The ordinary condenser departs considerably from such a performance. By considering the application of the integral equation for $\Delta\Theta$ to such a case it will be recognized that the relative space posi-

tions of the vectors representing the direct and indirect waves will be subjected to changes at every point where the slope of the frequency-time curve departs from a simple sine relation. The degree of distortion due to the presence of such irregularities may be considerable.

In Fig. 24 are shown some samples of "wobbled" carrier frequency records obtained at Stamford, Connecticut. For these records the

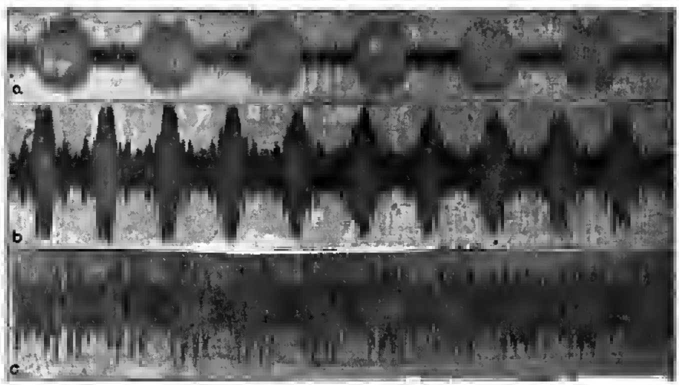

Fig. 24—Sample fast records showing distortion produced by intentional frequency modulation. *a* day record, *b* and *c* night records

carrier was wobbled at the rate of about 10 cycles per second. There is some uncertainty as to the range of frequency variation for these records although it was probably in the order of a few thousand cycles. By means of a constant frequency local oscillator the radio-frequency wave was stepped down in frequency to audio values which could be amplified and recorded.

The record (a) of Fig. 24 represents stable day-time reception. The record shows amplitude modulation due to the receiver characteristic alone. If the receiver were, as is desirable, capable of amplifying all the frequencies present in the received wave in the same ratio this record would be of constant width. In the subsequent examination of night records we must keep in mind the fact that the terminal apparatus is responsible for a certain part of the amplitude modulation. Its influence is readily recognizable.

The night-time records shown in (b) and (c) reveal a distinct distortion of the envelope aside from that present in the daytime record. Peaks appear and disappear within time intervals sometimes as short as a fraction of a second.

The record in Fig. 25 represents a slow picture of the changes shown in (b) and (c) of Fig. 24. If these wobbled frequency waves are studied carefully it will be noted that where a single peak stands at one moment there gradually comes in view another as if it were sliding from behind the first. The cycle length being about 1/10-second we may get some idea from this series of the rate at which the changes

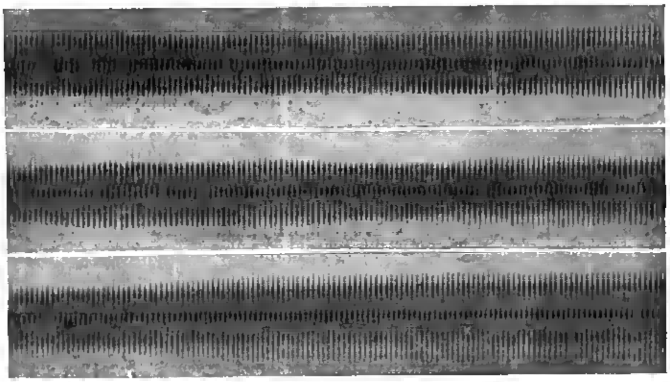

Fig. 25—Sample slow record showing distortion produced by intentional frequency modulation. Night record

take place. The presence of so many peaks in these records is attributed in part to the fact that the rotating condenser used gave a frequency change which was far from a simple sinusoidal relation.

Let us now return to the stepped-frequency method of obtaining the band fading pictures and ascertain why it has certain advantages. In (b) of Fig. 22 is shown the "equivalent" characteristic for the stepped condenser. During 1/2000 of a second (for the conditions so far assumed) in each step distortion may occur due to transient conditions, but during the remainder of the quarter second assigned to each step (for the records so far taken) a steady state is reached. Thus, theoretically, distortion occurs only during about 1/500 of the step interval. In (b) of Fig. 22 the lag is greatly exaggerated for purposes of illustration. This means simply that we have maintained constant frequency for a sufficient length of time to establish, before taking our picture, a fixed interference condition over the region including transmitter and receiver at least.

Daytime Field Strength Distribution

Thus far we have been dealing with the unstable phenomena of night-time transmission. Our interest has been directed almost

entirely toward variations with time. While the presence of wave interference has been detected, and the movement of this interference effect across the frequency band has been recorded, little effort has been made to form a picture of such interference in its space relation. A discussion of similar stable, daytime phenomena is therefore not out of place, and particularly so in view of an evident relation of the fickle nocturnal interference phenomena to the steady states which follow the appearance of daylight.

In a previously published map of field strength distribution in New York City,* it was indicated that the congestion of high buildings

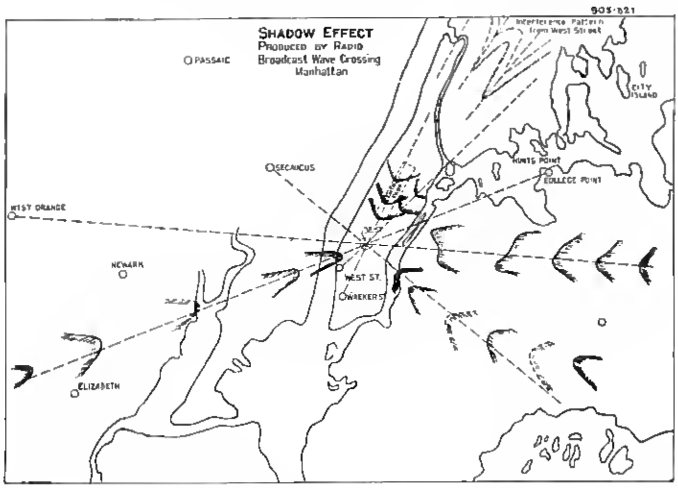

Fig. 26—Map showing location of radio obstruction on Manhattan Island as determined by the intersection of lines between various transmitting points and their corresponding shadows

just below Central Park cast a heavy shadow. More recently it has been determined from observations on a portable transmitter, set up at various points, that this building center is a consistent performer. The position of this obstruction is determined in Fig. 26 wherein only partial contours from maps for the indicated sites are given to prevent confusion. The intersection of these lines from transmitter to shadow, falls at approximately 38th Street in the vicinity of Sixth Avenue.

The dissipation of wave energy at such a point is probably the composite effect of many adjacent structures. Fig. 27 gives an elementary idea of how this can occur. The structures filling in

* See footnote 1.

each block are, of course, very well connected electrically by means of pipes, cables, etc., with those of adjacent blocks. Between each oscillating circuit (which is pictured as consisting of two buildings with earth connections) there exists a coupling which binds the whole system together more or less flexibly. Thus the obstacle offered by

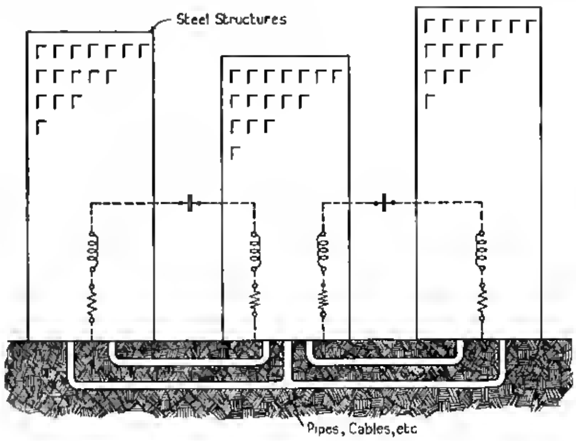

Fig. 27—Idealized picture of equivalent electrical circuit characteristics of high buildings

a group of buildings might be of a selective nature, and evidently its frequency characteristic may vary with direction.

Such an aggregate would, in addition to absorbing wave energy, produce a change in velocity or a refraction of the wave front. Some indication of such an effect will be discussed later. Before leaving the subject of shadows, however, let us get a physical picture of their significance.

From the transmitter a wave front expanding outward and upward encounters an obstruction which we shall assume is near the earth plane. The net result of this encounter is a weakening of the wave over an area near this plane, and probably a distortion of the energy-bearing fields. We might then imagine this shadow to be a tunnel-like region extending along the earth beyond the obstruction, and as having definite vertical as well as horizontal limits.

The aerial photograph of Manhattan and adjacent territory, shown on Fig. 28, will give a fairly clear idea of the conditions close to the transmitter. The major obstruction, the location of which has been previously described, is shown in its relation to the line of transmission toward the Riverhead and Stamford testing stations.

Fig. 28—Aerial photograph of Manhattan Island showing locations of transmitting station and obstructing high building area

Such barriers to wave travel, situated within a short distance from the source, seem, as we might expect, to have a more extensive and serious influence upon effective broadcast distribution than similar obstructions at greater distances.

It will be noticed that the obstruction falls very nearly upon the direct line from the transmitter to the Stamford testing station. This will also be evident later after an understanding of Fig. 29,

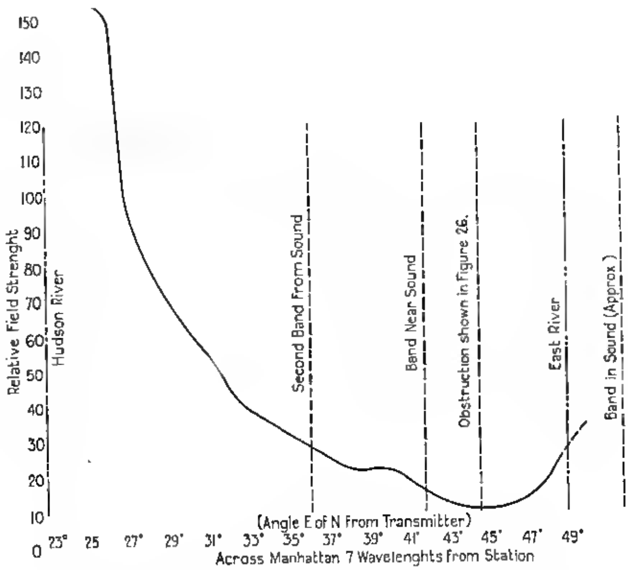

Fig. 29—Cross-section of radio shadow caused by high building area

wherein the position of the "Band Near Sound" represents also the bearing of the Stamford station. The Riverhead station is not directly in line with the major obstruction.

In certain sectors of the field strength contour map for station 2XB there appears to exist a kind of wavy displacement of the contour lines forming a partial pattern of peaks and depressions side by side. In general, this pattern must be differentiated from an ordinary shadow area. A remarkable example of this sort of field distribution is shown in Fig. 1 which is one section of a field strength survey made for station 2XB. These contours are based entirely upon daytime measurements, and represent a condition which is stable throughout the daylight period. Considerable difference in signal level is apparent within short distances across the direction of wave propagation. Two pronounced low signal channels extend ap-

proximately north-east across this region. These shift with change in frequency of the transmitted wave. Fig. 30 illustrates the space relations for such a movement. The full line curve shows a partial cross section of the contour map of Fig. 1 taken along a line approximately perpendicular to the direction of transmission 110 wave lengths from the transmitter. This represents relative field strength values for

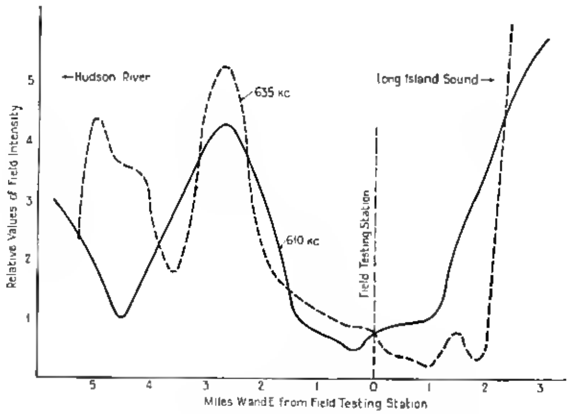

Fig 30—Cross-section of wave interference pattern showing change with frequency

610.0-kilocycle radiation. When the frequency is raised to 635.0 kilocycles, there occurs a movement of the peaks and depressions as is shown by the broken line of Fig. 30. Apparently the increased frequency causes these channels to be crowded together.

If we take sections of the field strength contour pattern in Fig. 1 and examine carefully the relative amplitude of peaks and depressions represented by these wavy lines we shall find that the ratio of field strength of the peaks to that in the depressions increases with distance from the transmitter. That is, the channels become more sharply defined as we move away from the transmitter. This ratio is shown approximately by the curves of Fig. 31. If these peaks or depressions were simple shadows they would maintain their relative values at a distance from the source or even tend to "heal" causing the ratio to fall rather than rise as is actually the case.

Within 14.4 wave-lengths (7.1 km.) of the transmitter the pattern, so apparent beyond 30 wave-lengths, merges into one deep shadow a cross-section of which is shown in Fig. 29. The abscissa of this curve is in degrees measured from the transmitter so that the center of the two most distinct low field strength channels extending north-

east may be inserted with their true radial relation. The two most evident in Fig. 1 are shown to be west of the line extending from transmitter through the center of the obstruction located in Fig. 26. The presence of Long Island Sound east of the geometrical center of the shadow has made an extensive survey of this section imprac-

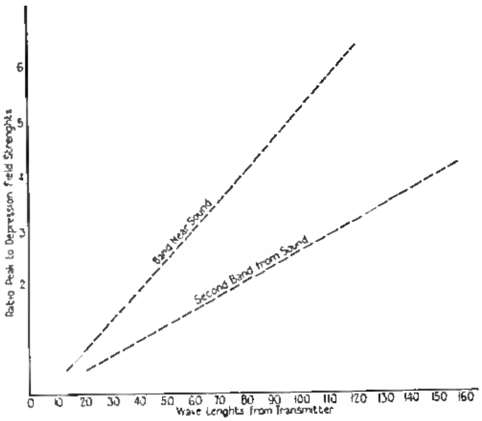

Fig. 31—Plot showing intensity of definition of wave interference pattern

tical. However, a single section taken across the Sound at about 90 wave-lengths from the station shows quite unquestionably the presence of a low channel about as indicated to the right of the obstruction designated in Fig. 29.

We have, therefore, a deep shadow with a more or less orderly array of maxima and minima within its limits. These maxima and minima grow more distinct at a distance from the transmitter, contrary to what we might expect for ordinary shadows. Furthermore, we find that they move as the frequency is changed. These facts lead to the belief that the phenomena in question are due to wave interference such as has already been described in connection with night-time fading, but characterized by very much smaller path differences. This daytime interference condition is fixed while we have seen that the nocturnal patterns appear to wander continually. To explain this more in detail let us return to the shadow and consider the phenomena which might accompany it in a little more detail.

The study of light has made available much information concerning the subject of wave interference. It is known, for instance, that the edges of shadows are not sharply discontinuous changes from light to darkness, but that a series of dark and light bands, called

diffraction fringes, are interposed between the full light and full dark areas. In our radio case the distance from the source to the obstruction and the dimensions of the obstruction are both very much smaller, in comparison with the wave length of the radiation, than for any ordinary case in light, but apparently the phenomenon is of the same general nature. By applying the ingenious principle of secondary sources used by Huyghens we might theoretically determine the distribution of the field beyond an obstruction placed in the path of the advancing radio waves. The basis of this principle is

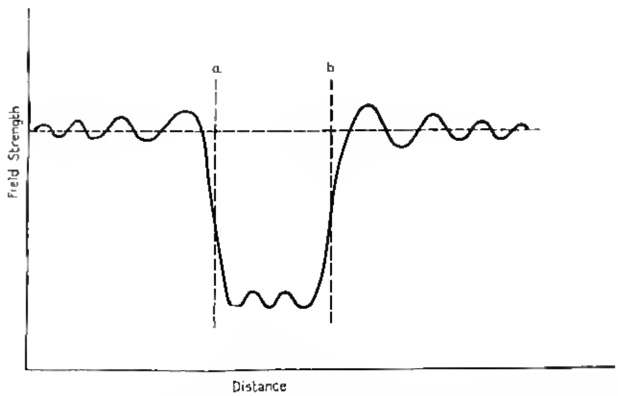

Fig. 32—Theoretical cross-section of radio shadow and associated wave interference pattern

the assumption that each elementary part of the advancing wave may be considered as a tiny transmitter. The effect at any point behind an obstruction, therefore, becomes the resultant effect, considering phase as well as amplitude, of the waves from all these miniature sources.

In Fig. 32 the region between vertical lines (a) and (b) represents the geometrical limits of the cross-section of a well defined shadow taken some distance behind the obstruction. An analysis of the resultant field using Huyghens' construction would show variations in intensity somewhat as represented by the full line. In other words the shadow will not be distinct but will have alternate maxima and minima within its geometrical limits and similar variations beyond the edges.

It is very likely, of course, that even in case the foregoing speculative analysis of the contour pattern extending north-east of 2XB is fundamentally correct, a great many other influences than that

of obstruction enter into the final field distribution. Relative attenuation of water and land appear to influence the distribution considerably though not as definitely as do steel structures close to the transmitter. Distinct minima appear both on the Hudson and on the Sound along radial lines extending from the transmitter.

Probably refraction of the wave front in passing across shore lines also enters into the shaping of this pattern.

Perhaps as good an elementary picture as any of the phenomena causing these patterns is that of a "dent" produced in the wave front by an encounter with a portion of New York City's impressive skyline. Since radio waves travel in a direction perpendicular to the plane containing the electric and magnetic fields, opposite sides of this "dent" would cross over one another with the result that an interference pattern would appear beyond the obstruction. An analogous situation exists when a water ripple passes a cluster of marsh grass which, damping its motion and retarding its progress causes part of the advancing front to converge and cross beyond the obstruction.

There is evidently a relation between day patterns such as have been discussed and night-time conditions. Just what this relation is offers some further opportunity for conjecture. In the first place quality distortion in transmission at night was, as previously explained, observed over parts of the region covered by the pattern shown in Fig. 1. The worst distortion seemed to be somewhat associated with the low field strength regions in this daylight survey. The distortion seemed also to be worse along the low channel extending in the direction of New Canaan, Conn., and beyond the 100-wave-length circle. It was particularly bad at a distance of some 140 wave-lengths from the station along this low channel where the field strength became so low in the daytime as to be unmeasurable with the set employed for the work. Accompanying the poor quality were fading and marked directional shifts.

Quality distortion though not so consistently severe at the Riverhead station as in the vicinity of Stamford was at times easily detectable by audible tests. Due to rapid attenuation of the radio waves traveling from the site of 2XB across Manhattan and the length of Long Island the field strength around Riverhead is generally low with higher levels north and south on the open waters of the Sound and Ocean respectively. Night-time fading at this point was representative of the variety which is usually found at distances of approximately one hundred miles from a broadcast transmitter.

The situation at Riverhead appears to be somewhat the same as

that which may exist over a large part of the broadcast area at a distance from the transmitter, while in the Westchester region we have an extreme and rather special circumstance. Field strength surveys have shown that there are indications of a daytime interference pattern over the Riverhead area but this pattern, such as it is, appears to be irregular and to lack the definition which makes the Westchester pattern so remarkable.

On the basis of the Westchester data alone we might build up a theory to the effect that night-time shifts of the stable daylight pattern were in some way responsible for quality distortion following the departure of daylight. Such a thought applied to the Riverhead case does not seem so reasonable since here the pattern is about one-quarter as distinct in terms of the ratio of maxima to minima values as the Westchester pattern. If, however, we presume that quality distortion may be expected in areas where daytime signals arrive *considerably attenuated* or so interfering as to simulate such an attenuated condition both situations are satisfied. After a consideration of the evidence at present available, such a conclusion seems attractive; that is, a daytime wave interference pattern alone is only an agency in night-time quality distortion in so far as its minima in combination with the general shadow effect are responsible for a low signal *directly* transmitted. Perhaps, in other words, the daytime field strength is a measure of *direct* night-time transmission, there existing in combination with this direct path at night a second, variable route of greater effective length. Probably close to the transmitter the "direct wave" is large compared to the "indirect" but shadows or interference may materially modify the ratio.

Night Distribution of Field Strength

By receiving simultaneously at several points the signal coming from a distant transmitter, it ought to be possible to detect the movement in space of these interference bands we have been discussing. The question immediately arises as to how far apart these distributed receivers can be placed without giving us an entirely discontinuous and misleading picture. For the first step toward recording space variations, in the vicinity of the Riverhead testing station, the receivers were spaced 1/16 wave length (30.5 meters), as illustrated in Fig. 33. It is necessary in making such determinations to transmit a single radio frequency, since we have already found that the interference bands for one component of a modulated wave are likely to be in a different position than those for another.

In order to receive and record the radio frequency wave it is, as has already been shown, convenient to use a local oscillator to beat it down to audible values. Since several oscillators for the separate sets are likely to produce mutual interference a common one was

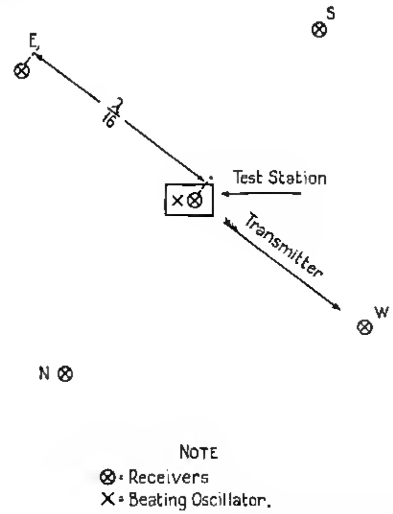

Fig. 33—Diagram showing space relation of receiving sets for special test

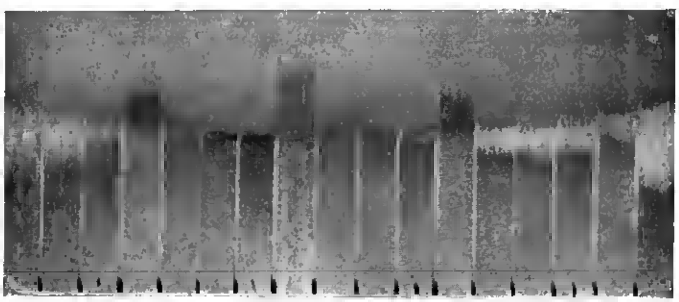

Fig. 34—Sample single-frequency fading record from spaced receiving sets

employed. This beating oscillator was situated at the testing station and the receiving antenna at this point was used as a radiator. In order to prevent overloading, the local receiver, the coupling to the receiver input coil was balanced to give a minimum of the local signal.

Fig. 34 is a sample of the record obtained. The continuous shadow band at the top represents the local receiver output. One oscillator

element was used for the other four receivers, their signals being recorded successively by a commutating device. Incidentally the interaction between these receivers was checked by observing the output of any one, while changes were made in the tuning of the others. The antenna was, however, so nearly aperiodic that no recognizable distortion or reradiation phenomena could be detected.

Fig. 35 illustrates compactly variations recorded by the oscillograph records (of which Fig. 34 is a sample), for a representative period of about five minutes. Even within the dimensions of 1/16

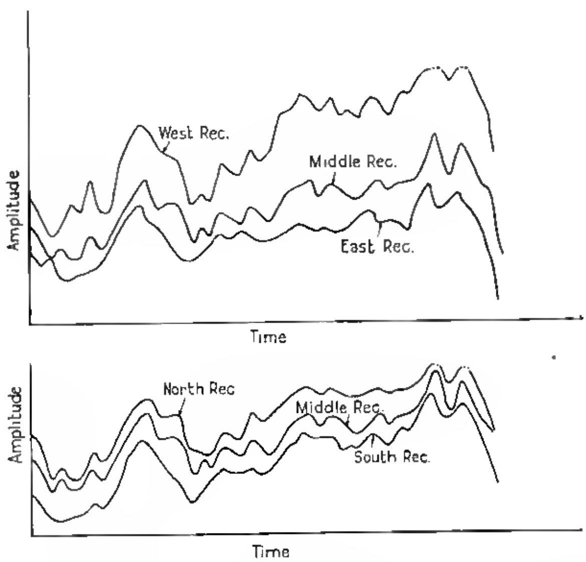

Fig. 35—Curves showing single-frequency fading on spaced receivers, condensed from long record

wave length there appears to exist transient field strength gradients in the direction of transmission. This is shown by a change in relative values, in the upper set of curves which represents field strength at points 1/16-wave length apart in the direction of transmission. The deviation is particularly noticeable in the relation between values for the local receiver and the "West receiver" which is in the direction of the transmitting station.

The lower set of curves, representing similar values across the line of transmission are much more nearly parallel. From the data so far obtained for the Riverhead testing site, it seems that transient nighttime field strength gradients are more generally evident in the direc-

tion of transmission than perpendicular to this direction. Upon these limited data one might be tempted to predict the presence of interference bands across the line of transmission.

The above discussion concerning space relation of field strengths has been included merely by way of contributing an additional bit of evidence to the theory that the erratic type of fading ordinarily

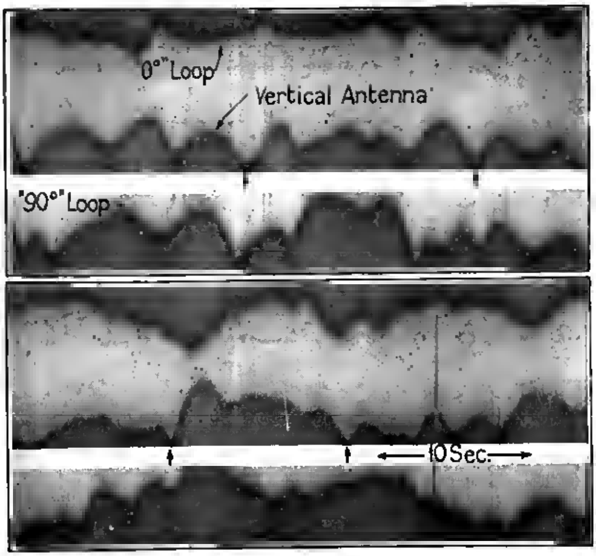

Fig. 36—Single-frequency fading record from vertical antenna and two-loop antenna crossed at right angles

experienced at night time is due to wave interference. The picture is very small in terms of wave lengths but considering its content, its very limits seem to imply wave interference rather than attenuation alone.

In connection with the wave interference theory thus far suggested as responsible for a major part of fading Fig. 36 is introduced as added evidence. The middle record of this group represents amplitude changes in the night-time reception of a carrier wave upon a vertical antenna. The upper and lower records represent the same for two loops turned at right angles to one another in the horizontal plane. By daytime tests the interaction of this combination was found to be negligible. Night-time fading recorded simultaneously for these three separate receivers occupying as nearly the same point in space as was possible, show that a high amplitude

signal may be coming in on both loops while the vertical antenna pick-up approaches zero. Several points of this kind are marked by arrows below the middle trace in Fig. 36.

There are at least two simple possibilities which might account for these relations. In case the wave approaches the receiving point from directly overhead, the vertical antenna would receive a "zero" signal while the loops would pick up an amount depending upon the state of polarization. If this be true, the records indicate a very rapid shift from the vertical direction of reception since the antenna minima are short lived most of them lasting at best a small fraction of a second.

On the basis of wave interference it is apparent that two waves approaching the receiving point in a 90-degree space phase relation and 180 degrees out-of-time phase could give a maximum signal on the two loops while that received on the vertical antenna was a minimum.

A compromise between these two viewpoints is probably a better guess than either one of them taken alone. That is, the existence of minima on the vertical antenna at the same moment that a strong signal is coming in on the loops is perhaps due to the interfering combination of waves having components in both the vertical and horizontal planes.

Quality Distortion

So far the data shown have been limited to the results of observations taken on special forms of transmission which are simplified for the purpose of clearly exposing the basic facts. We wish now to consider some of the more practical aspects of signal distortion. The first test which we made at our field test station was to record on slowly moving photographic paper tape and on the high speed film, the detected audio signal which resulted when the transmitter was modulated by a pure 264-cycle tone.

Fig. 37 is a sample of the general type of audio signal record obtained and Fig. 38 shows copies of the wave shape of the received signal, at particular times corresponding to the numbers of the oscillograms on the records in Fig. 37. The abrupt displacement of the timing trace indicates the point on the long record at which the snap-shot oscillogram was made. A peculiar characteristic of these records is the dark shadowy lines weaving back and forth through the band recording the complete signal. These dark lines correspond to the kinks in the wave shape shown in Fig. 38. As explained before, the darkening of the record is caused by the greater quantity of light

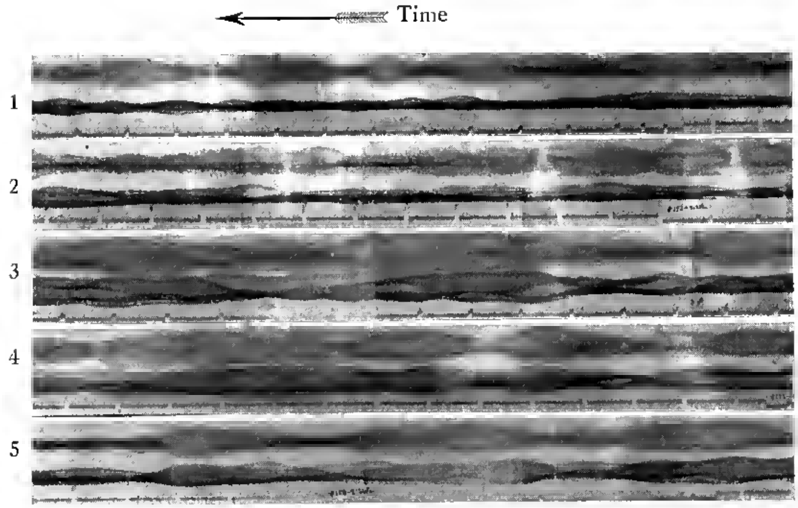

Fig. 37—Slow record of signal detected from tone modulated transmission showing the night-time distortion. Made at Stamford, Conn., May 15, 1924, 2:25 a.m. Upper trace signal from vertical antenna receiver and lower trace signal from loop antenna receiver, timing marks 2.6 seconds apart

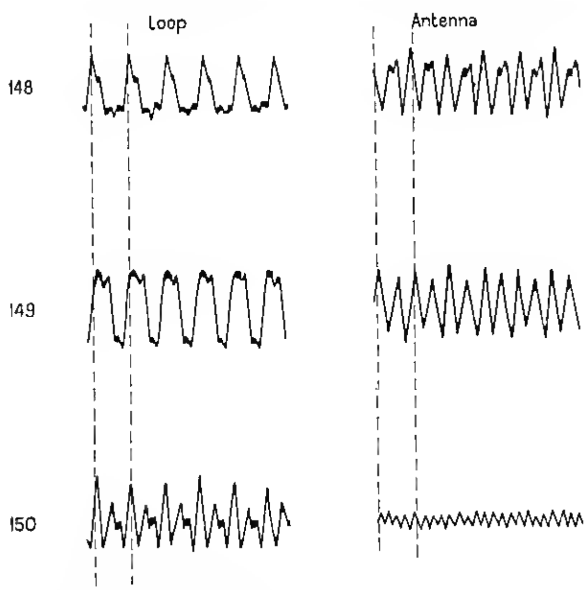

Fig. 38—Wave form of signals corresponding to numbered positions indicated on strips 2, 4, and 5, Fig. 37

affecting the record at these peak points. At the same time these observations were made, the wave shape of the signal rectified from the antenna current at the transmitter was recorded by an oscillograph. These oscillograms showed the signal to be free from distortion at the transmitter.

The weaving of these shadowy traces together with their width gives a record of the change in phase and amplitude of the irregularities in the wave shape of the signal. Although the wave shape of the signal is continually changing, it persists in substantially the same form for a great many cycles. Thus the record shows that, in the transmission of this simple tone modulated signal from the transmitting to the receiving antenna, it has been so modified that entirely new frequencies appear at the receiver. This receiver was shown by local tests to be free of any appreciable distortion within itself. While these new frequencies look like harmonics of the modulating tone in the snap-shot record it is obvious from the slow record that they are not true harmonics but that they differ from the harmonics by a very small amount and are incommensurable with the modulating tone since they undergo progressive but irregular phase changes with reference to it.

These records represent in a nutshell the signal distortion problem as it first presented itself to us. Our work then consisted in raveling out the complicated relations so that their nature could be ascertained and a theory of the causes established. In this paper, in the interest of clarity of presentation we have departed considerably from the actual order of the experimental work but at this point perhaps the actual order is best to follow for a moment.

With such a weird-looking distortion to analyze, and if possible eliminate, our first thought was as to whether the terminal apparatus might not involve unrecognized peculiarities which would be a contributing cause. Local tests and daytime tests of the receiving system absolved it from doubt and attention was focussed on the transmitting apparatus.

It was suspected that present day radio telephone transmitters leave something to be desired in regard to what we may call, for lack of a better term, their dynamic frequency stability. A very large percentage of the transmitters in use throughout the world today produce amplitude modulation of the carrier by the action of modulating tubes directly upon an oscillating tube circuit. It is to be expected that the cyclic changes in circuit conditions occurring at the modulating frequency will have some cyclic effect on the absolute frequency of the carrier and that this effect will be in the nature of a

wobbling or rapid shifting back and forth in frequency of the amplitude modulated carrier. In other words the carrier and side-bands, without change in their relative frequencies, would be subjected to "frequency modulation."

This sort of thing should be clearly distingushed from the slow wandering of frequency which, for instance, causes beat notes between carriers of different stations to drift gradually in pitch. What we have called "dynamic instability" is so rapid (being governed by the cyclic variations of the modulator) that it is difficult to observe by any aural method. Since the transmitter being used for our tests

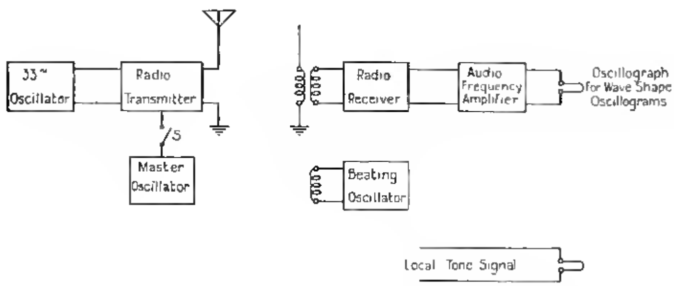

Fig. 39—Diagram of system used to measure frequency modulation

was a member of this almost universal class which employs modulating elements directly associated with the oscillator elements we determined to study this aspect of the transmission.

The following test was made to find out the extent of the frequency variation during the period of the modulating cycle. A schematic of the testing circuit arrangement is shown in Fig. 39. The plan was to modulate the carrier with 33 cycles, a tone so low in frequency that it would not be efficiently transmitted through the audio frequency amplifier connected to the output of the radio receiver. Then upon beating the received modulated carrier signal down to a frequency of about 1,000 cycles, an oscillogram of this signal would show a 1,000-cycle signal with a 33-cycle modulation in amplitude. Frequency modulation, if present, should then be easily discernible from the record. This experiment was made for daytime transmission and oscillograms (A) and (B) shown in Fig. 40 were obtained, one with the frequency of the beating oscillator greater than the carrier frequency, and the other with the beating oscillator frequency less than the carrier frequency. Both of these oscillograms show by the change in the frequency of the beat note signal

that frequency modulation occurs in the transmitter circuit. The frequency change is very apparent on the oscillograms when the lengths of one cycle at maximum and minimum amplitudes are compared. The reality of the effect is demonstrated in the two records, which by their difference show the reversal of the increased and decreased frequency points with reference to the modulation cycle when

A

B

C

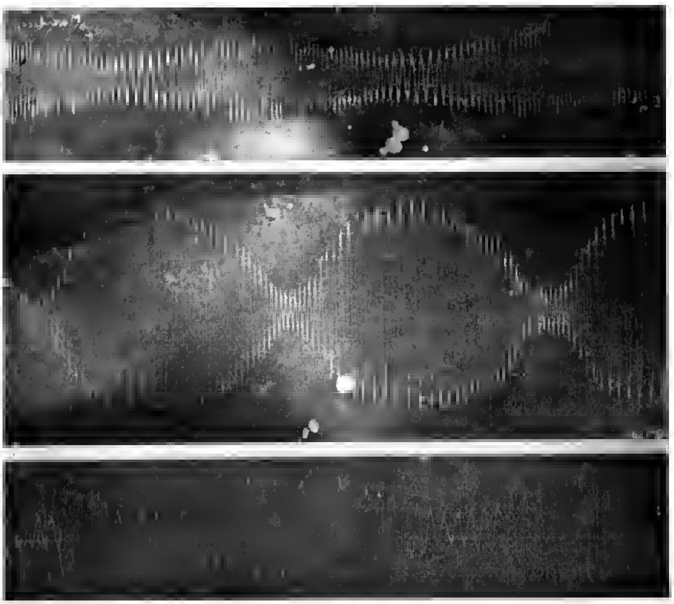

Fig. 40—Oscillograms showing frequency modulation accompanying amplitude modulation

the beating frequency is moved in frequency from one side of the carrier to the other.

The next step was to determine to what extent a stabilization of the carrier frequency to stop frequency modulation would affect the distortion of signals. True, master oscillator transmitters capable of giving the desired stability are not a new thing in the art. Several such transmitters were built by the Western Electric Company some years ago and used successfully in ship-to-shore radio telephone experiments[2] in which frequency stability was of considerable importance. To modify the ordinary broadcasting transmitter to in-

[2] See Fig. 1 and accompanying discussion in: Radio Extension of the Telephone System to Ships at Sea by H. W. Nichols and Lloyd Espenschied Proc. I. R. E., Vol. II No. 3.

clude this feature involves major mechanical changes and in order to provide a suitable arrangement for these tests the Bell Telephone Laboratories engineers merely added to the existing transmitter at station 2XB a temporary separate oscillator and high-frequency amplifier which could be connected to drive the oscillator tubes of the set as amplifiers. That this was free from frequency modulation is seen by comparing (C) of Fig. 40 with (A) and (B).

The transmission tests carried out with this arrangement yielded highly satisfactory results as indicated by a comparison of Fig. 41

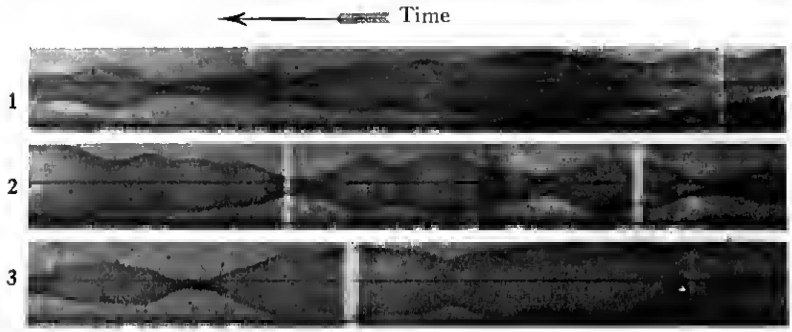

Fig. 41—Slow record of signal detected from tone modulated transmission with stabilized carrier showing reduction in distortion. Made at Stamford, Conn., Oct. 10, 1924, 3 a.m.

with Fig. 37. Fig. 41 like Fig. 37 is the detected result of a signal which started from the transmitter as a pure tone modulated signal, but it shows that much of the wave form distortion has disappeared, there remaining only a residuum which characteristically appears at the lower amplitudes of the signal. The probable cause of this residual effect will be discussed later. Tests of speech and music were concurrent with these findings. Using the normal transmitter, night-time transmission as received at the test stations was seriously distorted. When the stabilizing arrangement was employed this distortion was apparently eliminated except at the minima of fading.

Having arrived then at this practical result we wished to make further confirming tests, and tests to determine the whys and wherefores of the result. We have already detailed the more basic of these tests in previous sections of this paper and are now ready to consider the practical distortion records more carefully and to build up a theory to explain them.

The records shown in Fig. 42 are similar to the records in Fig. 37. They are shown here to illustrate the difference in the characteristics

of the wave form distortion variation that occurs from day to day. All these records were made at Stamford, Conn.

Strips 1 and 2—May 15, 1924—4:30 a.m.
Strips 3 and 4—Jan. 23, 1925—5:30 a.m.
Strips 5 and 6—Jan. 24, 1925—6:15 a.m.
Strips 7 and 8—Jan. 24, 1925—8:00 a.m.

There is a marked difference in the records obtained on January 23 and 24, which were made at the time an effort was being made to determine the effect of the solar eclipse on radio transmission. The peculiarly

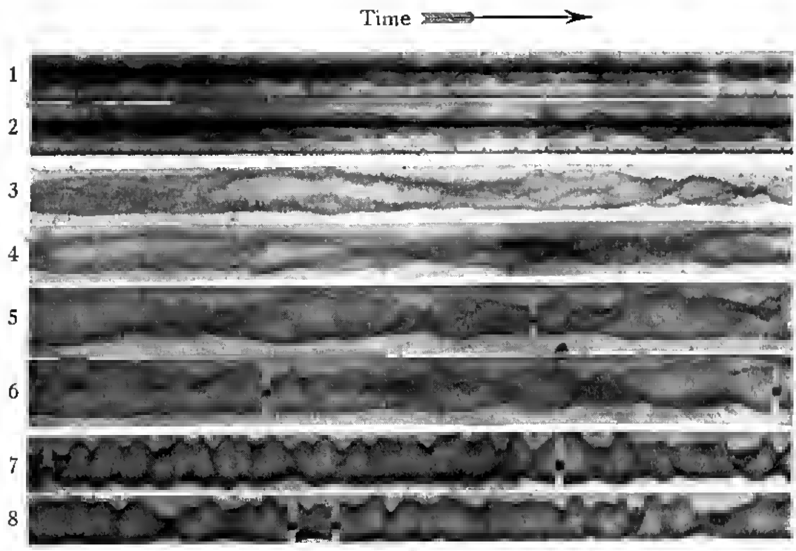

Fig. 42—Slow record of signal detected from tone modulated transmission taken on different days showing the changes in the character of the distortion

twisted appearance of the record obtained on January 24 is not very common in the records obtained. Most of the records have characteristics similar to those shown in Fig. 37. In the January 24 records there is a marked change in the characteristic configuration of the variation.

In order to obtain a record of the amount of wave form distortion resulting from frequency modulation present in the detected audio signal the circuit arrangement shown in Fig. 43 was used. This circuit was designed to analyze the wave form distortion when a 250-cycle

signal was used to modulate the carrier. Special precautions were taken to obtain a pure 250-cycle modulating tone. The wave shape of the signal detected from the carrier at the transmitter was frequently checked by observations with an oscillograph. The signals detected from the antenna current at the transmitter, both for the normal

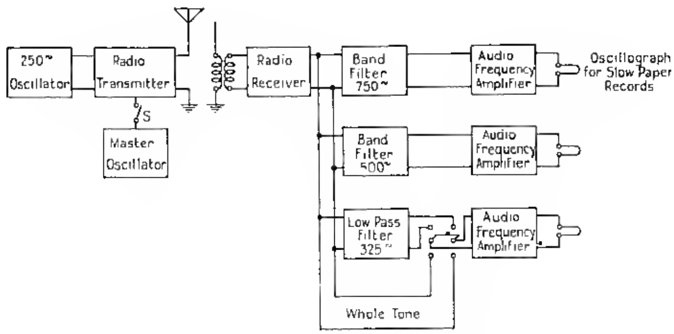

Fig. 43—Diagram of system used to obtain "Harmonic" analysis distortion records

transmitter with frequency modulation and for the stabilized carrier transmitter, were practically simple sine waves. The output circuit of the radio receiver was connected to a group of filters designed to transmit narrow bands of frequencies straddling the harmonics of 250-cycles.

While below we have referred to the frequencies passing these filters as "harmonics" it should be borne in mind that they are not necessarily *true* harmonics since they deviate very slightly from the true harmonic relation. The purpose of the test was to procure a record which would show at a glance the presence or absence of wave form distortion.

The input circuits of the filters were connected in parallel and the output circuits separately connected to the audio amplifiers arranged to operate the oscillograph elements. The input of one amplifier was arranged so that it could be switched either to the output of the filter passing 250-cycles or the output of the radio receiver. In this way a record could be obtained of either the whole tone from the receiver or only the 250-cycle component.

In Fig. 44, Strip 1 is a harmonic analysis record of the audio tone detected from the carrier and both side bands, transmitted with a stable carrier frequency. Strip 2 is a section of a record made a few minutes later when an unstabilized carrier was being used. On this record the lower trace is the 250-cycle component, the center trace the

500-cycle component, and the upper trace the 750-cycle component. The upper and lower traces have their zero lines at the edges of the strip. This record was made at Riverhead, L. I., April 30, 1925, at 3:33 a.m. Strip 2 is a section of a record made a few minutes later when an unstabilized carrier was being used.

The gain in the audio amplifiers connected to the outputs of the filters was adjusted to give nearly uniform transmission through the

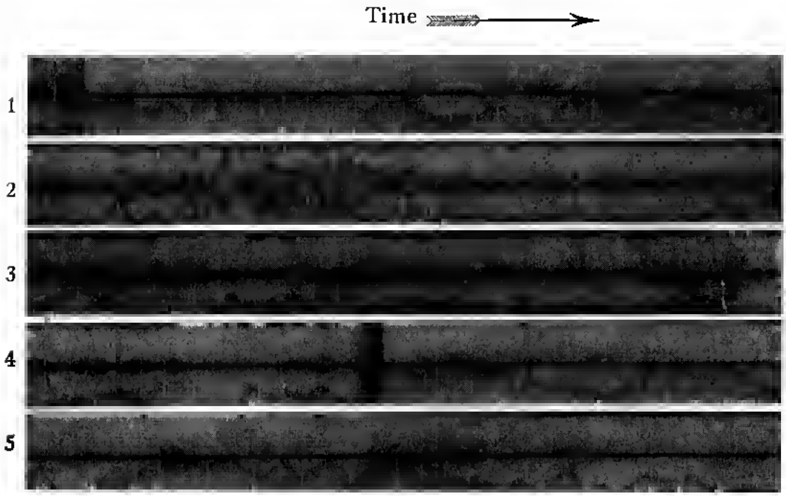

Fig. 44—Slow record made with system diagrammed in Fig. 43. Contrasting the distortion of detected tone transmitted by stabilized and unstabilized carrier frequency

receiving and recording apparatus for the frequencies recorded. Hence in these records the relative amplitudes of the fundamental and harmonics of the signal are directly comparable.

Strips 3, 4 and 5 in Fig. 44 are taken from a record made for the purpose of obtaining a comparison of the wave form distortion sustained by the detected audio signal transmitted by the normal transmitter with frequency modulation present and by a stable frequency transmitter. In each strip the lower trace is the whole tone from the output of the radio receiver, the middle trace the second harmonic (500 cycles) and the upper trace the third harmonic (750 cycles). Strip 3 and half of Strip 4 give the record obtained when the normal transmitter was used, and the remainder is the record obtained when the modified transmitter was used. There was a few minutes' difference in time between the ending of one transmitting condition to the beginning of the next during which the master oscillator control was switched

on at the transmitter. The receiving circuit was not changed during the making of this record, so that the results obtained from the two transmitters are directly comparable.

The record of the signal from the normal transmitter shows an abundance of second and third harmonics, at times equal in amplitude to that of the whole tone signal. The latter, of course, includes these harmonics. It will be noted also that dark line shadows run through the trace of the whole tone, indicating the presence of the wave form distortion. The signal from the stable frequency transmitter as shown by the record is practically free from wave form distortion. The trace of the whole tone is also free from any dark lines which would

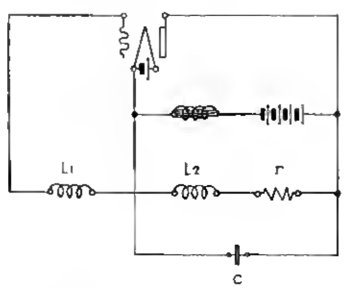

Fig. 45—Diagram of an oscillator circuit

indicate wave form distortion. This record is substantial evidence that a great deal of the wave form distortion may be eliminated when the carrier is stabilized. However, the selective fading still remains.

The selective fading we have already explained more or less satisfactorily and we find that it does not materially affect the wave form of audible frequencies transmitted by a modulated stabilized carrier unless its changes are more rapid than any we have recorded. The crippled state of originally perfect tone waves after they have been transmitted by an unstabilized carrier, we have just observed. Now let us consider the possible causes of this difference. The carrier stabilization referred to here, may we repeat, is not stabilization against slow variations in frequency from second to second or from hour to hour but rather against rapid variations within the cycle of the modulating frequency.

The reason for such changes over the modulating cycle is that the variation of the impedance of a vacuum tube across the oscillating circuit necessarily causes a variation in the nature period of the oscillation. As a simple case, the circuit in Fig. 45 is given.

H. J. Vander Bijl in his analysis[3] of this circuit gives the natural frequency of oscillation as

$$n = \frac{1}{2\pi}\sqrt{\frac{\left(1+\frac{r}{r_p}\right)}{L_2 C}} \tag{14}$$

when r_p is the plate resistance and the remaining constants are given in the illustration.

Direct modulation by the usual method involves a cyclic change in the value of plate resistance. Hence, according to the above equation, there results a cyclic change in frequency which, though relatively small, becomes of the utmost importance when subjected to the peculiar phenomena of night-time transmission.

By making certain assumptions concerning the nature of frequency variation as amplitude modulation takes place, it is possible to work out distorted waves corresponding to various assumed wave interference conditions at the receiver. Perhaps the most simple and instructive means for producing these distorted waves is by a graphical method.

The equation for modulation of a high-frequency wave by a single tone may be written

$$e = (A + kA \cos vt) \sin pt \tag{15}$$

When A represents the unmodulated amplitude of the wave, k is a factor determined by degree of modulation, v is an angular velocity of the tone wave and p is the angular velocity of the high-frequency wave. The amplitude factor in this equation may be considered as a vector which is undergoing a change in length in accordance with the term included in the brackets. For the purpose of our analysis we shall include the angular velocity imparted to this vector by the last term in the above equation, since we are interested in the envelope of the resultant high-frequency wave at the receiver and the relative phase relations for two waves directly and indirectly transmitted combining to form this resultant. Since both carrier waves are of the same mean frequency only the relative position need be considered.

Now in our graphical determinations for the case of two transmission paths different in length, we represent the two effective fields by vectors varying in length in accordance with the amplitude factor of equation (15). However, due to the difference in length of path,

[3] "Thermionic Vacuum Tube," by Van der Bijl, page 274.

the changes in length of one vector will *lag the changes in length of the other* by an amount

$$\phi = v\,(\Delta t) \tag{16}$$

when Δt equals the difference in time of transmission over the two paths and v is the angular velocity of the modulating tone. This angle ϕ for 500-cycle modulation may according to the data thus far described, amount to more than 90 degrees at the receiving points selected for observation.

In addition to the lag in amplitude there will be a lag in frequency change over the frequency modulation cycle. This lag which has

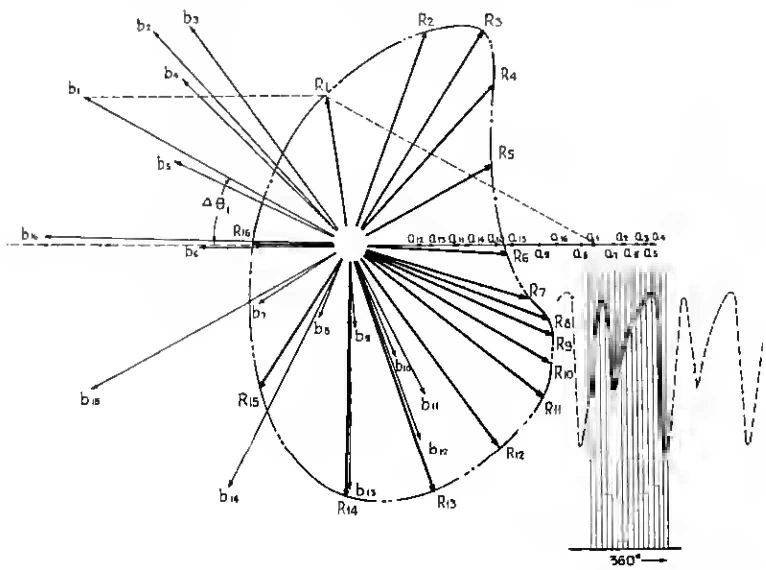

Fig. 46—Graphical method of synthesizing distorted wave forms caused by frequency modulation

already been shown in connection with the analysis of distortion in certain types of band fading records (see Fig. 22), becomes *a change in the relative phase angle* of the vectors under consideration. Thus our picture finally becomes one of two vectors changing in length, the changes in one continually lagging the changes in the other, the two vectors at the same time undergoing what we might term a relative angular wobble.

In Fig. 46 these relations are produced graphically. For our purposes we might assume that the vector representing one field is fixed and allow the other one to wobble the relative amount. At an

instant, for example, the directly transmitted field may he represented by a_1 in this figure. Assuming a difference in length of path, we may compute on the basis of the integral equation (13), the relative phase position of the vector representing the indirectly transmitted field b_1. The relative amplitude of this vector may also be determined by substituting $\Delta\phi$ in equation (15).

After establishing a sufficient number of vectors to represent the cyclic variation we may combine the respective components to obtain the resultant representative of the successive instants. These are

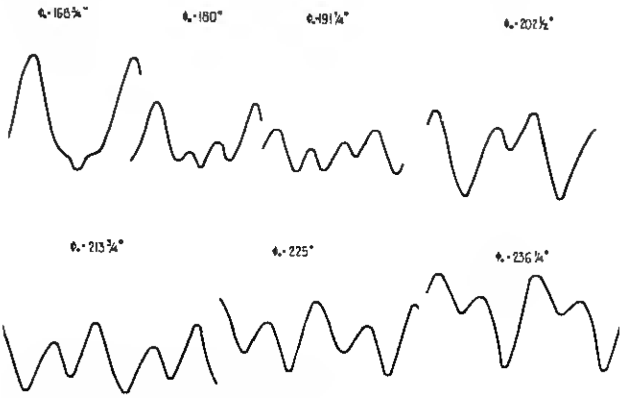

Fig. 47—Synthetic wave forms showing distortion due to frequency modulation

shown as R_1, R_2, R_3, etc., a broken line being drawn through their extremities to identify their positions. Now, if we plot these resultants as vertical ordinates in their successive time relation as shown on the lower right of Fig. 46, we have the envelope of the resultant wave at the receiver.

When the mean position of the two vectors (a) and (b) in Fig. 27 is 180 degrees separation, the signal is experiencing a fading minimum. When they are on the average in phase the amplitude is at a maximum. We can, therefore, trace a relation between quality distortion and fading by such an analysis, assuming a constant percentage modulation. Fig. 47 shows a series of high-frequency wave envelopes obtained by this method of graphic analysis. The mean vector relation is represented by ϕ_o, and for $\phi_o = 180$ degrees the fading may be considered at a minimum. The waves shown in Fig. 47 being envelopes of the high frequency will undergo certain changes in the process of detection. These, however, would only slightly modify the wave.

For purposes of comparison, a set of oscillograph pictures of representative received wave shapes is shown in Fig. 48. These represent the actual effect of night-time transmission with frequency modulation between 463 West Street, New York City and Stamford, Conn.; the modulating tone was a practically pure 264-cycle sinusoidal wave. The samples have been arranged in successive order

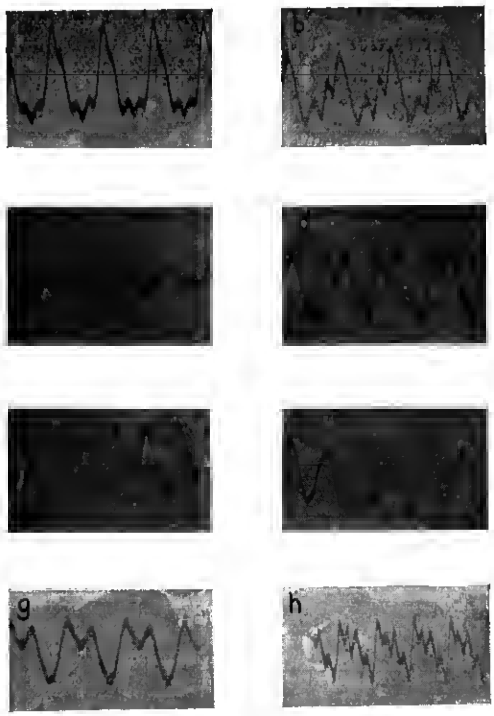

Fig. 48—Oscillograms showing actual wave forms with distortion resulting from frequency modulation

to correspond with the order shown in Fig. 47. There exists a striking similarity. Occasionally, however, the shapes predicted may depart considerably from those obtained experimentally. As an example of such a departure, the record (h) in Fig. 48 has been included. Such unusual samples may be due to a combination of waves arriving over more than two paths or it may be that the time variation of the frequency is far from the simple sinusoid which we have assumed. As a matter of fact, a critical mathematical treatment of this case shows that only an approximation of such a sinusoidal condition is

possible since as has been shown by Carson,[4] a frequency modulated wave of this character consists of an infinite series of fixed frequencies spaced at regular intervals either side of a "fundamental" carrier wave. Obviously only a small part of such a series could get out of the transmitter or into the receiver due to circuit selectivity. For the lower modulating frequencies, however, the approximation involved in the assumption of a simple sinusoidal variation is not far wrong since the amplitudes of these side frequency components fall off rapidly as their order in the series increases. While 150 wave lengths difference in path length has been assumed for the synthesis of the wave shapes in Fig. 47, this difference may according to the data obtained amount to much more than this.

It may well be asked why this frequency modulation, since it produces such marked distortion at night in certain places, does not also give rise to distortion by day or in locations where transmission is steady. A full answer to this question would be far from simple. But in brief it is because the carrier and side-bands shift in absolute frequency together as a unit so that their relative or difference frequencies which determine the audio signal remain unchanged. Another way to put it is that the detector operates on the envelope of the high-frequency signals and is blind to the frequencies contained within the envelope except insofar as they affect the latter. However, since frequency modulation appreciably widens the frequency band occupied by the radio signals it is to be expected that the tuned circuits in the receiver would have some reaction on those louder portions of the signal for which the amplitude modulation and therefore the frequency modulation is large. The perfection with which broadcast signals may be received under suitable conditions leads one to believe that this effect must be small.

Fading in Relation to Form of Transmission

It has been shown that serious wave form distortion of the reproduced signal may result if frequency modulation occurs with the amplitude modulation and the transmission is subjected to night-time conditions. This distortion from frequency modulation can be eliminated by stabilizing the carrier frequency. There remain some wave form distortion and the annoying amplitude changes caused by selective fading which is one of the most serious present day problems in radio transmission. Let us now consider the nature and cause of this

[4] See "Notes on the Theory of Modulation," by John R. Carson, Proc. Institute of Radio Engineers, February, 1922.

residual wave form distortion and some further consequences of selective fading under the assumption that there is no frequency modulation involved.

The process of detecting audio signals from radio frequency signals is, at least in its simpler aspects, well understood, but it may not be generally appreciated that the action is such that the detected signals may be greatly modified by changes in the relative amplitudes and phases of the carrier and side-band components such as may result from their transmission through the medium. That the amplitudes and phases of the carrier and side-band signals are not necessarily received in the same relation that existed as they left the transmitter has been pointed out earlier, in the discussion on selective fading.

The usual expression for a high-frequency carrier wave of frequency $p/2\pi$ modulated by a low-frequency wave of frequency $v/2\pi$ is

$$e=A[1+a\sin(vt+\phi)]\sin pt$$

where A is the carrier amplitude, a, the percentage modulation and ϕ the starting phase of the modulating tone with reference to the carrier. Expanded into its components this becomes

$$e=\frac{A_1a}{2}\cos(pt+vt+\phi_1) \qquad \text{(the upper side band)}$$

$$-\frac{A_2a}{2}\cos(pt-vt-\phi_2) \qquad \text{(the lower side band)}$$

$$+A_3\sin pt \qquad \text{(the carrier)}$$

where $\phi_1=\phi_2=\phi$ and $A_1=A_2=A_3=A$ as the waves leave the transmitting antenna.

In the receiving set this function is squared by the action of the detector and, neglecting direct currents and frequencies above the audio range, the result is

$$\frac{a}{2}A_3[A_1\sin(vt+\phi_1)+A_2\sin(v+\phi_2)]+A_1A_2\frac{a^2}{4}\cos(2vt+\phi_1+\phi_2) \qquad (17)$$

of which the first term represents the fundamental frequency of the original modulating tone and the second term the second harmonic.

From this expression several conclusions can be immediately drawn. Due to the action of the detector there is always some slight wave form distortion as is evidenced by the presence in relatively small amplitude of the second harmonic. In the ordinary case this is negligible. The first term contains the carrier amplitude as a

factor but the second term does not. Thus, if selective fading erases the carrier at any time, reducing its amplitude to zero or a small value, the signal, represented by the fundamental tone, practically disappears, *even though the side-bands have not faded out*, and there remains only the harmonic. This is the residual distortion shown in Fig. 41 and which can often be heard during a fading out period. It is caused by the two side-bands beating together in the detector. We have here exposed a fundamental defect in the usual form of modulated signal transmission. The amplitude of the received signal is subject to all the whims of the carrier and to paraphrase freely an old saying we might remark that a signal is no stronger than its carrier. We may at once conclude that one way to reduce fading is to suppress the carrier and resupply a constant amplitude carrier at the receiving station.

Analyzing further the first term of the expression representing the detected signal, the first part of the bracketed portion results from beating together in the detector of the carrier and upper side-band and the second part from the carrier and lower side-band. It is clear that one of the side-bands may fade out completely and the other will still bring in the signal, provided the carrier is not also lost, with a phase shift to be sure but nevertheless not seriously reduced in amplitude. In telephony this kind of phase shift is relatively unimportant. Here we have an evident advantage in transmitting both side-bands since they support each other's frailties. But if the two side-bands suffer phase shifts in transmission, as we have earlier shown may be produced by wave interference, such that ϕ_1 and ϕ_2 differ by π radians or 180 degrees, the two components will cancel each other provided their amplitudes A_1 and A_2 remain equal. In other words all three components—carrier and both side-bands—may arrive at the receiver with full amplitude and yet no signal will be detected from them except a second harmonic component. This is obviously a disadvantage of transmitting both side-bands since, at such an instant, if one of them were eliminated the signal would reappear.

We conclude that there is, on the basis of such a brief analysis, not much to choose between single side-band and double side-band transmission when the carrier is transmitted also.

But if we wish to realize the advantages of carrier suppression a choice is not difficult. A carrier suppression system in which both side-bands are transmitted requires that the replacement of the carrier at the receiving station be done with almost absolute accuracy as to frequency and phase, a thing which involves very serious prac-

tical problems. On the other hand if but a single side-band is transmitted the difficulty is reduced to placing the carrier within a very few cycles of its correct position. The allowable departure will depend on a number of things but there is reason to believe that for high quality transmission it must be very small, perhaps no greater than two or three cycles.

With the single side-band carrier suppression method, invented by John R. Carson, the radiation is stripped down to the minimum which will fully transmit the telephonic signals and this reduces to a minimum the exposure of the signals to the ravages of selective fading. If the spacing interval of the fading is relatively narrow as in the cases we have examined hereinbefore, this form of transmission would not fade seriously in average volume but would be subjected to a continual changing of its frequency-amplitude characteristic, that is to say individual frequency components would fade progressively as the minima of the selective fading wandered back and forth across the frequency range encompassed by the single side-band. If the spacing interval of the fading were very large so that the minima were very broad of if some other, at present unexplored form of fading which covers a wide band at one time were acting, the signal would fade in average volume but the range of its variation would be only the square root of that of a carrier transmitted signal, since only the side-band would fade and the locally supplied carrier would remain unchanged.

The extent to which these theoretically drawn conclusions may be realized in practical application is yet to be determined but we have a few records bearing upon the matter which at least do not run contrary to them.

All of the transmission tests where the radio signal was beat with a local oscillator and the detected beat note observed, were equivalent to single frequency single side-band transmission with carrier suppression, the local oscillator functioning as the carrier suppressed at the transmitter. In this case, for which a number of records have already been shown, the detected signal is in proportion to the product of the amplitudes of beating oscillator and received radio signal. The phase of either does not affect the amplitude of the audio signal. Hence, the only important modification of the original signal is the variation in the amplitude resulting from selective fading.

Unfortunately we have no records in which a direct comparison is made between single side-band transmission with and without carrier suppression but the case can be visualized from the record shown in Fig. 12 or 13. Here each one of the frequencies recorded may be looked upon as a single side-band frequency which has been detected through

the agency of the resupplied carrier of the beating oscillator used to bring them down to audio-frequency. If now we were to take two of these frequencies shown on the record and multiply their amplitudes together at each point we would obtain the amplitude of the signal which would result if one of them were a single side-band and the other its accompanying carrier. It is obvious that the fading variations would thereby be increased in amplitude and rapidity.

In order to obtain a comprehensive picture of the relative advantages of radio transmission using a carrier and one side-band as compared

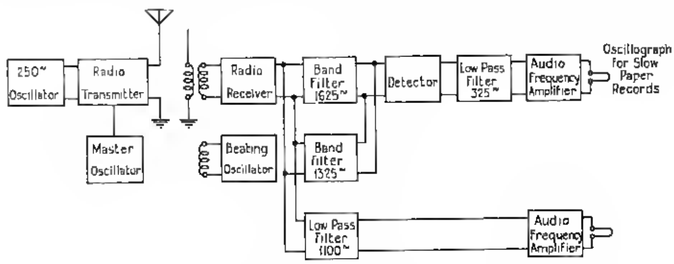

Fig. 49—Diagram of system used to obtain records of transmission with carrier and one side band and carrier and both side-bands

with the common practice of transmitting both side-bands, the following tests were made. The schematic diagram of the circuit arrangement is shown in Fig. 49. At the transmitter the carrier and both side-bands are transmitted and at the receiver they were selected out by means of filters in the manner previously explained. The signals from the filters corresponding to the carrier and lower side-band were applied to the input of a detector circuit and from its output the detected difference signal was selected by a low-pass filter. This signal was equivalent to that which would be received if only the carrier and one side-band were transmitted. From the output of the radio receiver a branch circuit goes to a low-pass filter which transmits only the signal detected from the carrier and both side bands, suppressing from this circuit the higher frequency signals corresponding to carrier and side-bands produced by the beating oscillator and received signals.

By making simultaneously a record of these two signals a direct comparison is obtained of the effect of selective fading on their amplitudes. Fig. 50 shows samples of several such records made at Riverhead, L. I. The modulating frequency for strips 1, 2 and 3 is 250-cycles, and for strips 4 and 5, 500-cycles. The record on strip 3 is shown on account of the peculiar characteristic of the signal fading, for considerable periods of time remaining at relatively low amplitude.

In these oscillograms the upper trace is the record of the signal from the carrier and both side-bands, and the lower trace the signal from the carrier and lower side-band.

These records illustrate by giving a graphic comparison the effect of the phase changes of the component signals in the case where the

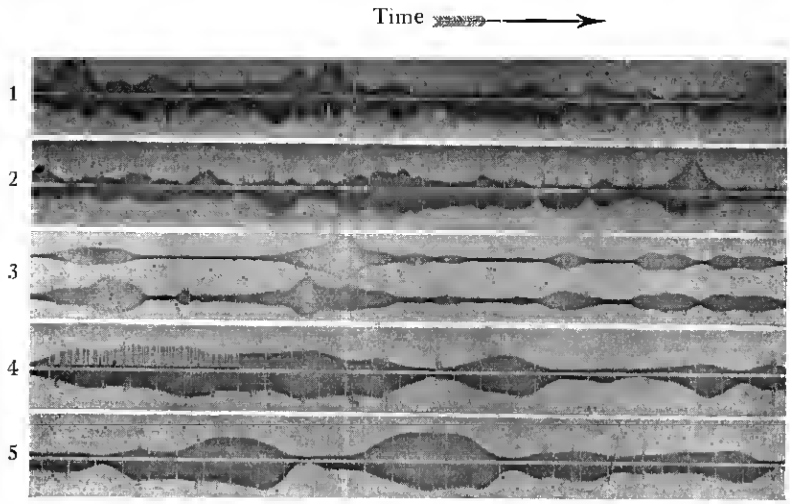

Fig. 50—Slow record comparing the signal detected from carrier and one side-band with signal detected from carrier and both side-bands. Made at Riverhead, L. I. Upper trace carrier + both side-bands, lower trace carrier + one side-band. Strips 1 and 2, July 22, 1925, 1:46 a.m. 250-cycle modulating tone. Strip 3, July 21, 1925, 3:10 a.m. 250-cycle modulating tone. Strips 4 and 5. July 23, 1925. 2:47 a.m., 500-cycle modulating tone

signal is detected from both side-bands. The amplitude of the signal from both side-bands in some instances is very small but appreciable amplitude is still indicated at the same instant for the signal from one side-band. This is explained as meaning that the side-band phases were such as to make the component signals 180 degrees out of phase after detection and that the amplitudes of the components were practically equal. The reverse situation is also observed where the amplitude of the signal detected from the lower side-band is zero and appreciable signal is recorded for the case where both side-bands are used. This is interpreted to mean that the side-band signal was eliminated by selective fading. In this event it was, of course, not contributing to the signal which was detected from both side-band signals. The recorded signal comes from the other side-band which evidently was not eliminated at that instant by selective fading.

Visual observations made with the cathode ray oscillograph, which unfortunately furnishes no permanent record of transient effects, confirmed the strip records in regard to the reality of there being side-band phase variations. From equation (17), it is seen that if these variations occur the fundamental of the detected tone signal at the receiver will not bear a fixed phase relation to that detected from the transmitting antenna current while if there are no such changes the phase between these two tones would remain constant. The locally detected tone and the tone detected from the transmitting antenna current and brought to the receiving station over telephone wires, were applied to the two pairs of deflecting plates in the cathode ray oscillograph. Since the deflections caused by these two pairs of plates are at right angles to each other the resulting Lissajous figure from two sine waves of the same frequency will be a slanting line, an ellipse or a circle depending on their phase and amplitude relation. The actual figures were observed to change progressively through this range of shapes, the changes following roughly the magnitude and rapidity of the fading. The effect of amplitude changes on such figures is quite distinct from the effect of phase changes and there was no difficulty in separating out the evidence of large phase changes.

Considering only the above theories and facts there appears to be a reasonable basis for a conclusion that the best form of radio transmission for use in broadcasting is single side-band with carrier suppression. But on practical grounds we do not believe such a conclusion is justified. The fading and distortions which we have made much of in the preceding pages are not experienced by the majority of broadcast listeners when they listen to local stations. To require these listeners to provide themselves with more complicated and expensive receivers, simply to allow more distant or less favorably situated listeners to obtain better reception, seems neither reasonable nor desirable. The art offers several other possible avenues toward improvement much less difficult of application and it must be remembered that radio broadcasting is already reaching a degree of standardization and a volume of existing receiving equipment which rules that changes must come slowly and without serious prejudice to the existing order.

Conclusions

Subject to the limitations imposed by the scope of our investigations the following conclusions may be drawn:

Fading can be quite sharply selective as to frequency and the evidence points toward wave interference as the cause.

The evidence for wave interference indicates that some of the energy of received signals reaches its destinations by a circuitous route and suggests that this route is by way of upper atmospheric regions.

Quality distortion may result from dynamic instability of the transmitter.

Fixed wave interference patterns in connection with shadows sometimes exist in daytime transmission.

CORRECTION SLIP FOR ISSUE OF JANUARY, 1926

Page 172: Equation should read

$$\Delta d = \sqrt{y^2 + 4h^2} - y.$$

Page 177: Equations (9), (10), (11) and (13) should read

$$\Theta_1 = 2\pi \int_0^t [F_o + f \sin r\,(t - d_1/V)]\,dt, \tag{9}$$

$$\Theta_2 = 2\pi \int_0^t [F_o + f \sin r\,(t - d_2/V)]\,dt. \tag{10}$$

$$\Delta\Theta = \Theta_1 - \Theta_2 = 2\pi \int_0^t F_o\,dt + 2\pi \int_0^t f \sin r\,(t - d_1/V)\,dt$$
$$- 2\pi \int_0^t F_o\,dt - 2\pi \int_0^t f \sin r\,(t - d_2/V)\,dt, \tag{11}$$

$$\Delta\Theta = \frac{2\pi f}{r}\,[(\cos rt - 1)\,(\cos r\,d_2/V - \cos r\,d_1/V)$$
$$+ \sin rt\,(\sin r\,d_2/V - \sin r\,d_1/V)]. \tag{13}$$

Delete (12)